

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2002, 09:44:53 ; Search time 86.21 seconds
(without alignments)
4590.135 Million cell updates/sec

Title: US-09-817-538-2

Perfect score: 1611
Sequence: 1 atgtctggggctcttcgccg.....tccctcacgtttcttccc 1611

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:**
- 2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:**
- 3: /cgn2_6/ptodata/2/lna/6A.COMB.seq:**
- 4: /cgn2_6/ptodata/2/lna/6B.COMB.seq:**
- 5: /cgn2_6/ptodata/2/lna/PCRTUS.COMB.seq:**
- 6: /cgn2_6/ptodata/2/lna/Backfiles1.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1498	93.0	2111 1	US-08-528-255A-2
2	1498	93.0	2111 1	US-08-717-365-2
3	502.6	31.2	1943 4	US-09-282-305-7
4	501.8	31.1	2019 4	US-09-282-305-5
5	460.8	28.6	1576 4	US-09-282-305-9
6	459	28.5	1826 4	US-09-282-305-1
7	392.4	24.4	600 4	US-09-385-982-402
8	329.2	20.4	1475 4	US-09-282-305-3
9	204.6	12.7	841 4	US-08-998-416-317
10	77.2	4.8	7218 1	US-08-232-463-14
11	50	3.1	3489 2	US-08-728-323A-1
12	50	3.1	32207 2	US-08-770-379-20
13	50	3.1	32207 4	US-08-757-669A-20
14	50	3.1	32207 4	US-09-230-371A-20
15	48.4	3.0	390 4	US-09-197-6649-7
16	44.2	2.7	1931 2	US-09-130-114-2
17	43.8	2.7	3223 1	US-07-980-528-1
18	43.8	2.7	6755 3	US-08-931-999-4
19	42	2.6	480 2	US-08-726-306A-185
20	41.8	2.6	5433 3	US-08-929-329-1
21	41.6	2.6	1042 4	US-09-276-531-77
22	41.6	2.6	43795 3	US-08-742-185-101
23	41.2	2.6	2233 1	US-08-145-705A-1
24	40.8	2.5	49136 4	US-09-422-869-1
25	40.8	2.5	51259 3	US-08-781-891-209
26	40.6	2.5	289 4	US-09-007-005-17
27	40.6	2.5	289 4	US-09-244-796-17

28	39.6	2.5	15378 3	US-08-785-420-1	Sequence 1, Appli
c 29	39.2	2.4	16442 3	US-08-781-891-208	Sequence 208, App
30	38.4	2.4	2277 1	US-08-676-967-2	Sequence 2, Appli
31	38.4	2.4	2277 2	US-08-676-974-2	Sequence 2, Appli
32	38.4	2.4	2277 2	US-09-098-487-2	Sequence 4, Appli
33	38.2	2.4	2223 1	US-08-257-073-4	Sequence 4, Appli
34	38	2.4	1298 3	US-08-948-703-3	Sequence 3, Appli
c 35	37.8	2.3	2188 4	US-07-865-662F-10	Sequence 10, Appli
c 36	37.6	2.3	43676 3	US-08-374-219B-10	Sequence 10, Appli
37	37.6	2.3	80246 4	US-09-356-952-12	Sequence 12, Appli
38	37.6	2.3	80595 4	US-09-078-294-3	Sequence 4, Appli
39	37.4	2.3	301 2	US-08-948-703-3	Sequence 3, Appli
40	36.8	2.3	3211 2	US-08-574-959A-8	Sequence 23, Appli
41	36.8	2.3	3211 4	US-09-357-014-8	Sequence 8, Appli
42	36.8	2.3	3901 4	US-08-574-959A-6	Sequence 8, Appli
43	36.8	2.3	3901 4	US-09-357-014-6	Sequence 6, Appli
44	36.6	2.3	1436 1	US-08-106-981-5	Sequence 6, Appli
45	36.6	2.3	1436 1	US-08-106-981-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-528-255A-2
Sequence 2, Appli
Patent No. 5659016
GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: FURUKAWA, YOICHI
TITLE OF INVENTION: RPD PROTEIN AND DNA
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: FLYNN, THIEL, BOETTLE & TANNIS, P.C.
STREET: 2026 Rambling Road
CITY: Kalamaazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb storage
OPERATING SYSTEM: IBM PC/XT/AT Compatible
SOFTWARE: WordPerfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/528,255A
FILING DATE: September 14, 1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP6-227876
FILING DATE: 22-SEPTEMBER-1994
APPLICATION NUMBER: JP7-183763
FILING DATE: 20-JULY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Terryence F. Chapman
REGISTRATION NUMBER: 32549
REFERENCE/DOCKET NUMBER: Furuya Case 1335
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2111
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
LIBRARY: human fetal lung CDNA library
FEATURE:

CURRENT APPLICATION NUMBER: US/09/282,305
CURRENT FILING DATE: 1999-03-31
PRIOR APPLICATION NUMBER: 60/080,563
PRIOR FILING DATE: 1998-04-03
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 7
LENGTH: 1943
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (57)..(1610)
US-09-282-305-7

Query Match 31.2%; Score 502.6; DB 4; Length 1943;
Best Local Similarity 65.5%; Pred. No. 2,6e-143;
Matches 768; Conservative 0; Mismatches 399; Indels 6; Gaps 2;

QY 134 gaggaagctgttactactacgagcggtatgtgaaattactattatgagcaagcca 193
DB 122 gggcgcggtgtctactacttaccgacccggtatgtggcaactactactacggcgagcca 181
QY 194 cccaatgaagcctcacccgaatccgcatgactataattgtctgctcaactatgtctcta 253
DB 182 tcggatgaagcgcgacgcgacatccgcatgagcactgcgtcgtcgcgctacgacctct 241
QY 254 ccgaataatggaatctatcgcctccacaaacgaatgcttgaggagatgacaaagtaacca 313
DB 242 caacagatgagcggtgtacccgccaacccgcccgcgcgcgcgcgcgcgcgcgcgcgcgc 301
QY 314 cagcagatgactataatctctggtcctccatccgcagataacatgtcggagtagacag 373
DB 302 cgcgcagcagctacatcaacttctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 361
QY 374 caagcagatgacagagattcaacgttggtagagactgtccagatctcagatgacctgttga 433
DB 362 ccgctgtcgtcgaagcgtctcaacgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 421
QY 434 gttctgtcagttgttctactggtgtgtctgtgtgcaagtgctgtggaactaataagacaga 493
DB 422 ctctgcgcagactatgctggcgccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 481
QY 494 gacggagcagcgtggaattggtcgtggggcctgcgacacatgcaagaagtcgcgagcgc 553
DB 482 ---tgacatcgcaatcaactggtcgggggcctgcgacacgcaagaagtgcgagcgctc 538
QY 554 tggcttctgtacgtcaatgatatgctctggtccatccctggaactcgttaagtatacca 613
DB 539 gggcttctgtacgtcaatgatatgctctggtccatccctggaactcgttaagtatacca 598
QY 614 gaggggtgtcatgacatgatatatcaatgattgagcggtgaggaagcgcttcta 673
DB 599 gagagttctgtatgcgatatcgatatccacatgtagatgagggagggcttctca 658
QY 674 caacacggacggcgtcatgactgttcttcaataagtaaggagagttacttccaggaac 733
DB 659 caacacagtaggtatgactgtctgtccacaagttgtgtatatttcccaaggaac 718
QY 734 tggggaactacgggataccggggtctggcaagaacagtatttgtgttaactaccgct 793
DB 719 aggggatalccgtgacatgtagcactcaaaaggaagtaacttccctgaatgtccctct 778
QY 794 ccgagacgggattgtatgagagtlccatagagcaatttcaagccggtatgtccaaat 853
DB 779 agatgattgagattgagagaaagcaacagtccttlltaagcaatcatgagcaaat 838
QY 854 aatgagatgttccagcagtagtgcgtgttcaacagtgtggtcagactcccatctatcg 913
DB 839 tatgtaggtttccgcccgtgtgcaggttgtgtctcaagttgtgtcgtatcttctgtctcg 898
QY 914 ggatcggttaggtgtgtcaatctatcaaaagacagcgaagtggtggaatttgt 973

DB 899 ggataggttgggtgtcttcaactctcaatcaaaagtcacgtgaggtatgttaggtatat 958
QY 974 caagagcttaaccctgcctatgtatgtctggaggggtgtttacacatctgtgaagt 1033
DB 959 gaggcttctcaacgttccatgttgccttctgtgtgtgtgatatatacaagaatgt 1018
QY 1034 tggccggtgtgacatagagacagcgtgtgcccctgtgatacagagatccccaagagct 1093
DB 1019 tgcagcgtgtgtgttatagagcgtgaggtgtcttggccacagctgtgaagaagaat 1078
QY 1094 tccatagatgactacttgaacttggaccagatttcaagctccacatcagctcttc 1153
DB 1079 gctgttaatgagtactaagatgactcgtgtccagattacactcttcaatgtgtccacaag 1138
QY 1154 caatagactaacccgaacacgaatgagtagctgtgagaaagatcaaacagcgactgttga 1213
DB 1139 taacatggaacaaaataacacgaacaaatgtagatatacagatcttaacttcttga 1198
QY 1214 gacattagaatgtgcgcgacacgtgggttccaacgcgagcga--ttcctgaaga 1270
DB 1199 taatcttcaaaacttgcagatgctcctcagttgtccacttcaagagagatctcctgaac 1258
QY 1271 cgcacatccctggagagagtgatgagtagca 1303
DB 1259 agaatacctgagcaagatgaaatcaagatga 1291

RESULT 4
US-09-282-305-5
Sequence 5, Application US/09282305
Patent No. 6287843
GENERAL INFORMATION:
APPLICANT: Baldwin, Donald A.
APPLICANT: Briggs, Steven P.
APPLICANT: Crane, Virginia C.
TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
FILE REFERENCE: 5718-44,
CURRENT APPLICATION NUMBER: US/09/282,305
CURRENT FILING DATE: 1999-03-31
PRIOR APPLICATION NUMBER: 60/080,563
PRIOR FILING DATE: 1998-04-03
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 2019
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (140)..(1459)
US-09-282-305-5

Query Match 31.1%; Score 501.8; DB 4; Length 2019;
Best Local Similarity 65.3%; Pred. No. 4.7e-143;
Matches 753; Conservative 0; Mismatches 397; Indels 3; Gaps 1;

QY 121 cgcagggcaccggggaagctgttactactacgagcggtatgtgaaattactatt 180
DB 189 cggagcgggtcggaagcgccgctgtctacttctacgacgcggaagtgaggcaactactct 248
QY 181 atggaagaagccacccaatgaaagcctcaacgcgaatccgcatgactatattgtctgtcca 240
DB 249 acgggacagggccaccccgatgaaagccgcacgcatcgcgatgacccaacgcgtcgtccggcc 308
QY 241 actatgtcttaccgaaatgaaatctatcgcccttacaagaacaaatgtctgagga 300
DB 309 gctacgacctctcgtacacagatgacagtggttccgcctcaacctgtccgcgacgcgagacc 368
QY 301 tgaccaagacacagcgatgactataatcttctggtcctccatccctgcctcaagtaaca 360
DB 369 tctgcgccttccacgcgcgacgattacgtctcttctcctcgtcgttcaaccccgaaagc 428

Oy	361	tgctcgaggtacagaagacagatgacagagattcaacgttggtagagactgtccagattccg	420
Db	429	agcagcgaccagatctccgcgcgtccaaagcgtctcaacgtccgacgtcgagagactgtccccgtcttcg	488
Oy	421	atggcgcgtttgaaattctgtccagttctctctctcgtgtgttcctctgttgagaaatgtctgtgaac	480
Db	489	acggtctctacagtcttctgtccagacgtacgcggggggtcctgtttggcggcgcggtccaaagc	548
Oy	481	ttaaagacagcagacagagacatccgcgttgaaattgggcttgvgggggtccgtcacatcgaaga	540
Db	549	tcaaccatgagcca---tgaatattcgccatcaactggtcgcggtcgagatccacacagccaga	605
Oy	541	agtcgcggagcatctcggtcttctgttaagtcgaatgatattctgttcggccatcccgagacgc	600
Db	606	agttgcgaggtccctccgggttttctgttaagttaatgatactgtccctcgccatccctccgacgtcc	665
Oy	601	taaaagtatcacacagaggtgtctgtataattgacataattacataatccaatgtgtgacgcgcg	660
Db	666	tcaagatatacaacagcgcgtcttcgttaagcttgagacttgatatccacacagggagcgcggtcg	725
Oy	661	aagagcctctctacacacagacgcgggtgcatagtactgtctcttcaataagttagagagt	720
Db	726	agagaggtcttttataccacagacgcgggtgtagtcagcttccatccacaaagttttagaagt	785
Oy	721	acttcccaaggaactggtggagactcaggtgatalacggggctgtgcagaagaagatttatgtg	780
Db	786	attctccctggagacggggagacattcgtgattgttggtaactcaaaagggtataattatctccc	845
Oy	781	ttaaactacccgtctccgagacgggatttgatgacagctcctataagggccattttcaacgcgg	840
Db	846	tgaattgtccctctcgagcagatgatttgatgtagagagctacacagtcgtttgttcaagccaa	905
Oy	841	tcatgtccaaagttaaavgagatgttccagcctgavgcgttggtcttcaacatgvtggtccag	900
Db	906	taatggtgcacaggtgtatgtgaggttcttcaacccctgtgcagtcgtgtccacagtgtgtgtcg	965
Oy	901	actccctatccggggagatcgggttgattgtgtcttaatatctatctcaaaaggacacggcaagt	960
Db	966	attcaattgtcgggtgtgacagagtgtgggtgttcttcaacctctctataagggtccagcagaagt	1025
Oy	961	gtgtggaattgttcaagaagctttaacctgcctatgctgtgaatgtcttgvggagcgatgtgtaca	1020
Db	1026	gtgtgaagatttatgaggttcccttcaacgtcccgctgtgtctgtgtgtgtgtgtgtgtata	1085
Oy	1021	ccattgttaacgtttgcccgggtgctctgacatataavgacagcvtgtggccctgtacagsga	1080
Db	1086	ccataaagaacagttgtccgcggtgtgtgtgtgtcagcagacgaagattgcgcctgtgtcatgagc	1145
Oy	1081	tccttatgtgcttccatataactgaactactttaactttgagccaaagtttcaagctcc	1140
Db	1146	tcaactgacagaatgaccctcaactgaagtactatgaaattttgtgttccagattcaccctcac	1205
Oy	1141	acatcagctccttccaatatagtactaacacagacacgcgaatgagatcctgtgaggaatcaaac	1200
Db	1206	atgtcgtctccaagaatacatgttgagataataaaacacacgcgataatttgatgtgataaac	1265
Oy	1201	agcgacgttttgagaacctagaatgtctgcgcgacgacacctgtgggttccaaagcagcgca	1260
Db	1266	caaaactctagataactcttcaanaactccgacatgctctcctagtgttcagtttcaagac	1325
Oy	1261	ttcctgaggaagcg	1273
Db	1326	gaactcctgaggc	1338

[illegible]


```

RESULT      7
US-09-385-982-402
; Sequence 402, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCNDA-260XX
CURRENT APPLICATION NUMBER: US/09/385,982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 402
LENGTH: 600
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(600)
OTHER INFORMATION: n = A,T,C or G
US-09-385-982-402

Query Match          24.4%; Score 392.4; DB 4; Length 600;
Matches Local Similarly 98.1%; Pred. No. 5.5e-110;
Match 407; Conservative 0; Mismatches 7; Indels 1; Gaps 1

OY   1183 accctgagaagatcacaacagcgactgtttgtagaaccttgaattgtctgcgcagcaccty 1242
Db    1 acctggagaagatcaaacacagcgacctgtttgtagaaccttgaattgtctgcgcagcaccty 60

OY   1243 gggtcacaacgagcaggcgattcccttgagagaccatcccttgagagagtgg -cgatgagac 1301
Db    61 gggtccaatatgaggcgattcccttgagagaccatcccttgagagagtggcgatgagac 120

OY   1302 gaagcgcgccctgaaaagagcatctcgattctctctctcgaaacgaattcctgttg 1361
Db    121 gaagcgcaccctgaaaagagcatctcgattctctctctcgaaacgaattcctgttg 180

OY   1362 gaaggtctctcgattctgtaagagggagggggggcgccgaagaaccttccaacttc 1421
Db    181 gaaggtctctcgattctgtaagagggagggggggcgccgaagaaccttccaacttc 240

OY   1422 aaaaaagccaagaggttcaaaaacagagatgtaaaaagagaagaccagaagagaagaaa 1481
Db    241 aaaaaagccaagaggttcaaaaacagagatgtaaaaagagaagaccagaagagaagaaa 300

OY   1482 ggaatcaccgaagagaagaaaaccaaagagaggaagccgaagcccaaggggtccaagag 1541
Db    301 ggaatcaccgaagagaggaagaaaaccaaagagaggaagccgaagcccaaggggtccaagag 360

OY   1542 gaggcacaagtgcctctaagtacaccttcacgctctgactcttcctgctgatacct 1596
Db    361 gaggcacaagtgcctctaagtacaccttcacgctctgactcttcctgctgatacct 415

RESULT      8
US-09-282-305-3
; Sequence 3, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: MAIZE HISTONE DEACETYLASES AND THEIR USES
; FILE REFERENCE: 5718-44,
```

CURRENT APPLICATION NUMBER: US-09/282, 305									
CURRENT FILING DATE: 1999-03-31									
PRIOR APPLICATION NUMBER: 60/080,563									
PRIOR FILING DATE: 1998 -04 -03									
NUMBER OF SEQ ID NOS: 18									
SOFTWARE: PatentIn Ver. 2.0									
SEQ ID NO 3									
LENGTH: 1475									
TYPE: DNA									
ORGANISM: Zea mays									
FEATURE:									
NAME/KEY: CDS									
LOCATION: (29)..(1084)									
US-09-282-305-3									
Query Match									
Best Local Similarity 20.4%; Score 329.2; DB 4; Length 1475;									
Matches 499; Conservative 0; Mismatches 283; Indels 0; Gaps 0;									
QY	528	cacccctgcaaaagaatccgagcgcatctggtcctctcttgaagtaagatgcgtctggcc	587						
DB	164	caccgcgcgaagaagcgagagcccggtctctgtaagtaacgaactcgtctccgc	223						
QY	588	atccctggaactgctaaagatcacacagagggtgctgtatcatgtgacattgacaat	647						
DB	224	atcccgagctccctaagttcacacagcgctgtgctatagtagacattgatgtccacac	283						
QY	648	ggtgacgacgcttggaagagcgctcttcaacacagagacggcggtgtatgactgtcccttca	707						
DB	284	ggagatgacgctggaagagagcgctcttccatcacaaacgagatcagatgtctcccttcc	343						
QY	708	aagtaatgagagtaattccacagaactcgggagacttaacggagataccgggcttgcaagac	767						
DB	344	aagtaatgagagatcttccctcgtgactgtgacatacactgcagcttgaggcagccgaagg	403						
QY	768	aagttatgctgtttaactaacccgctcggagacgggattgtagaagagatcctatgagcc	827						
DB	404	aagcatltagctctgtaatgtctccctcgagtgatgatacgaagacccacttcgttgt	463						
QY	828	atttcaagccgctatgctccaaagtaatgagatgttccacgctatgctggtgctta	887						
DB	464	ctgttcatatgatacatataagaagaattatgagaggtttatcacgacgagctggttgcctc	523						
QY	888	cagtgctgcagactcccatctcgggagtcggtaggttgcttccaatctatctaca	947						
DB	524	caatcggaagcgtacactcttggctggagacaggttagtgcttccactgctctgtaag	583						
QY	948	ggaacgcgcaaatgtgtygtaatttgcgaagacttaaccctgcctatgttgctgggga	1007						
DB	584	gtctatgtcatgtgcctccgttccctttagtgcgtaaatgttccctatgttggttaagt	643						
QY	1008	ggcgtgtgttaacacctcgtlaacgtltggccggtgtgtagacatatagacaagctgtggcc	1067						
DB	644	ggtgagaggttaacacatcagaatgtltgaagcgtgtgtgtcgaagacgcgaagtgtc	703						
QY	1068	ctgtagtacggagatcccttaatgagcttccatcacatagactacttgaatcttggaca	1127						
DB	704	gttgaagttagaacctgataacaagctgtccttcaaatgattacatagaaacttggccct	763						
QY	1128	gattccaagctcacatcagtcctcttccaatagactaacacagaaacgcgaatgagtcctg	1187						
DB	764	gattatctcttcatatccaccaaaagtglttgaaaacctgtgaatccacaagaagcttg	823						
QY	1188	gagaagatcaaacagcgactgtlttgaagaaccttagaayctgcgcgacgaacctgtggtc	1247						
DB	824	gagaacataaagaacatgatatgttgaagaacctgtcaaaagatagaaactgttccacgaact	883						
QY	1248	caaaagcgagcttccgttgagagcgcaccccttggaagagatgycgatbtagaagcaagac	1307						
DB	884	caattccatgacagacgcttcagaccttgagacttccagagagaaaagaggaacatgagac	943						
QY	1308	ga 1309							

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic acid search, using sw model

Run on: June 8, 2002, 09:32:12 ; Search time 2132.14 Seconds
(without alignments)
10198.015 Million cell updates/sec

Title: US-09-817-538-2

Perfect score: 1611
Sequence: 1 atgtctgggggtctctgccgcg.....tccctcagcttcttccccc 1611

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 segs, 6748477542 residues

Total number of hits satisfying Chosen parameters: 27472414

Minimum DB seq length: 0
Minimum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	942.8	58.5	999 9	AL518877 AL518877
2	897.8	55.7	919 9	AL531607 AL531607
3	876	54.4	910 9	AL546378 AL546378
4	847.6	52.6	989 9	AL541430 AL541430
5	847.4	52.6	981 9	AL559849 AL559849
6	839	52.1	847 9	AL558916 AL558916
7	824	51.1	832 9	AL522933 AL522933
8	797.4	49.5	1627 11	BC013141 Homo sapi
9	789.2	49.0	954 10	BC054715 602714475
10	779.4	48.4	910 10	BG756668 602715527
11	778	48.3	982 10	BG756668 602715527
12	737.4	45.8	778 10	BM477303 AGENSCOURT
13	736.8	45.7	1007 10	BM452740 603053373
14	732.4	45.5	803 10	BM452740 AGENSCOURT
15	728.6	45.2	1034 10	BM472180 602743069
16	726	45.1	865 10	BM472180 AGENSCOURT
17	725.8	45.1	854 10	BG753951 602709473
				BM761949 603048783

18	725.6	45.0	748 10	BG821396	BG821396 602724847
19	725.6	45.0	794 9	AU141862	AU141862 AU141862
20	724.4	45.0	1014 9	BB611086	BB611086 BB611086
21	718	44.6	1190 10	BG387283	BG387283 602456030
22	717.4	44.5	871 10	BE747628	BE747628 601575753
23	714.2	44.3	743 10	BT906302	BT906302 603063251
24	712	44.2	734 10	BT830892	BT830892 603081018
25	708.2	44.0	770 10	BT832602	BT832602 603081964
26	706.6	43.9	905 10	BT855025	BT855025 603382030
27	703.6	43.7	1078 10	BE794277	BE794277 601591572
28	696.2	43.2	703 9	AL546172	AL546172 AL546172
29	692.4	43.0	734 9	BG739817	BG739817 602630342
30	692.2	43.0	834 10	BE742797	BE742797 601574414
31	683.6	42.4	947 10	BG386766	BG386766 602455037
32	683	42.4	32	BE568103	BE568103 602183902
33	681.8	42.3	841 10	BG468565	BG468565 602510349
34	681.2	42.3	701 10	BG248617	BG248617 602400789
35	673.2	41.8	1026 10	BT766905	BT766905 603053149
36	672.6	41.8	914 10	BG248511	BG248511 602400669
37	669.8	41.6	988 10	BG031523	BG031523 602299352
38	666.2	41.4	697 10	BT260793	BT260793 602970769
39	665.8	41.3	878 10	BE560409	BE560409 601347136
40	664	41.2	817 10	BE559940	BE559940 601345219
41	663	41.2	860 10	BE559855	BE559855 601345754
42	662.8	41.1	916 10	BE743289	BE743289 601574890
43	658.6	40.9	955 10	BE796991	BE796991 601588015
44	657.4	40.8	777 10	BT869316	BT869316 603396215
45	657	40.8	849 10	BE747533	BE747533 601577745

ALIGNMENTS

RESULT 1
LOCUS AL518877 999 bp mRNA linear EST 13-FEB-2001
DEFINITION AL518877 LTR_NFL011.NBC1 Homo sapiens cDNA clone CS0DA011YG11 5
PRIMER mRNA sequence.
ACCESSION AL518877.1 GI:12782370
VERSION AL518877.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 999)
AUTHORS Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1. 999
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DA011YG11"
/clone_lib="LTR_NFL011.NBC1"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library is not normalized, but is the control for the normalized libraries. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville , Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

Db 841 TCCCTGACAAAGATTCCTGTGAGGAGAGTTCTCCGATTCGAGAGAGAGAGAG 900
 QY 1395 gggggccgca 1404
 Db 901 GGGGGCGGMA 910

RESULT 4
 AL541430
 LOCUS AL541430 898 bp mRNA linear EST 16-FEB-2001
 DEFINITION AL541430 LTI_FL002_P11 Homo sapiens cDNA clone CS0DE006Y103 5 prime
 , mRNA sequence.
 ACCESSION AL541430
 VERSION AL541430.1 GI:12872494
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 898)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polajcs, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 source
 1. 898
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DE006Y103"
 /clone_lib="LTI_FL002_P11"
 /lab_host="DH10B"
 /note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
 cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-stranded cDNA was digested with Not I
 and cloned into the Not I and Eco RV sites of the
 pCMVSPORT 6 vector. Library was constructed by Life
 Technologies. Contact : Feng Liang Life Technologies, a
 division of Invitrogen 9800 Medical Center Drive Rockville
 , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
 fliang@lifestech.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 227 a 204 c 237 g 224 t 6 others
 ORIGIN

Query Match 52.6%; Score 847.6; DB 9; Length 898;
 Best Local Similarity 98.1%; Pred. No. 6,4e-166;
 Matches 883; Conservative 5; Mismatches 9; Indels 3; Gaps 3;

QY 104 gggcaagatggcgcaagcagcagcggcgaggaagccttactactacgagggga 163
 Db 1 GAGCAAGATGGCGCACACCCAGGAGAGAGTCTGTACTACTACACGGGA 60

QY 164 tgttgaatactactatgaacaagcccaatgaagcctacacgaatccgatagc 223
 Db 61 TCTTGAATTAATTAATTAATGAAGGACCAACCAATGAAGCCTACCGATGAC 120

QY 224 tcataattgtctgaactatgctctacgcgaataatgaaatctatccctcaca 283
 Db 121 TCATATATTTGCTGCTCAATGCTCTACGAAATAATGAAATCTATCGCCCTCACA 180

QY 284 agccaatgttgagagatagcaagagagagagagagagagagagagagagag 343
 Db 181 AGCCAAATGCTGAGAGATAGCAAGTACCAAGCGATGACTATTAATTTCTTGCGCTC 240

QY 344 catcgctcagataacatgtcgagagagagagagagagagagagagagagag 403
 Db 241 CATCGCTCAGATTAATGAGAGAGAGAGAGAGAGAGAGAGAGATTCACAGCTTGTA 300

QY 404 ggaactgcagatattcgatgcctgttgagctgtcagttgtcactgtgtgtcgt 463

Db 301 GGACGTCCAGATTCAGATGAGCCTGTTGAGTTCTGCTAGTTACTAGTGAGTTCT 360
 QY 464 ggaagtgctgtgaaactaataagagagagagagagagagagagagagagag 523
 Db 361 GGCAAGTGTGTGAACCTTAATTAAGCAGAGAGAGAGAGAGAGAGAGAGAGAG 420

QY 524 cctgacacatgcaaaagctcgaagcctcgtctgtctgtcgtcgtcgtcgtcgtc 583
 Db 421 CCTGCACCATGCAAGAAATCCGAGCATCTGCTTCTTCTTACGTCAATGATGCT 480

QY 584 ggcacatcctggaactgtaaaatataccagagagagagagagagagagagag 643
 Db 481 GGCCATCTCGAAGACTCTTAAGATATCAGAGAGAGAGAGAGAGAGAGAGAG 540

QY 644 ccatgtgagagcgtgtggaagagcctcctcaacacagagagagagagagagag 703
 Db 541 CCATGCTGACGGCGTGTGAAGAGGCGCTTACACACAGAGAGAGAGAGAGAG 600

QY 704 tcataagtag 762
 Db 601 TCATAGTATGAGAGATTAATTCACAGAGAGAGAGAGAGAGAGAGAGAGAG 660

QY 763 aagacaagtattatgctgttgaactacccgctccgagagagagagagagagag 822
 Db 661 AAGGCAGATTAATGATGCTGTAACTACCCCTCGAGACCGGATGTGATGACGAGTCTATG 720

QY 823 aggcatttcgaagcgtgcatgctcgaagagagagagagagagagagagagag 882
 Db 721 AGCCATTTTCAGACCGGCGCTTAAAGATTAATGAGATTTTCCAGCCTATGCGGCTG 779

QY 883 tctacagatgtgctcgaactccctactcgtggagagagagagagagagagagag 942
 Db 780 TCTTACAGTGTGCTGAGACTCCCTATCTGGGAGATGAGTATGCTTCAATCTAATA 839

QY 943 tcaagagagagcgaagtgctgtggaatgttgaagagagagagagagagagag 1002
 Db 840 TCMAAGGACACCGCAAGTGTGATTTGTCAAGAGCTTTAA-CTGCCTATGCTGATGC 898

RESULT 5
 AL559849
 LOCUS AL559849 981 bp mRNA linear EST 16-FEB-2001
 DEFINITION AL559849 LTI_FL011_BCI Homo sapiens cDNA clone CS0DG005Y621 5 prime
 , mRNA sequence.
 ACCESSION AL559849
 VERSION AL559849.1 GI:12905737
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 981)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polajcs, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
 source
 1. 981
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DG005Y621"
 /clone_lib="LTI_FL011_BCI"
 /sex="male"
 /tissue_type="B cells from Burkitt lymphoma"
 /lab_host="DH10B"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-stranded cDNA was digested with Not I and Eco
 into the Not I and Eco RV sites of the pCMVSPORT 6 vector.

Library was constructed by Life Technologies. Contact :
Feng Liang Life Technologies, a division of Invitrogen
9800 Medical Center Drive Rockville, Maryland 20850, USA
Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com>

BASE COUNT 271 a 226 c 267 g 213 t 4 others
ORIGIN

Query Match 52.6%; Score 847.4; DB 9; Length 981;
Best Local Similarity 97.5%; Pred. No. 7.2e-166;
Matches 900; Conservative 3; Mismatches 14; Indels 6; Gaps 4;

```
QY 643 accatggtacagcgtgtagaaggccttaccacacgagccgggtcactgttctct 702
DB 1 ACCATGTTACGCGCTGGAGAGGCGCTTTACACACGAGACGGGGTCATGACTGTCTCT 60
QY 703 ttcataagatggaagctctccacgagacttggagactacaggaatccggagctgga 762
DB 61 TTCATAGATGAGAGTACTTCCAGAACT-GGAGACTACGGGATACGGGGCTGGCA 119
QY 763 aagacaagtattatgctgttaactaccgctccgagacggatgtagaagatccatg 822
DB 120 AAGGCAAGATATATGCTGTTAACTACCGCTCCGAGACGGGATGATGAGAGTCTATG 179
QY 823 aggcactttcaagccggtcactgctcaaatgtagatgttccagctagtggctgg 882
DB 180 AGGCCATTTTCAAGCGCGTATGCTCCAAAGTATGAGATGTTCCAGCTAGCGGTGG 239
QY 883 tcttaagtggtgctcagactccctactcgggagatcggttagtgcttcaatctcta 942
DB 240 TCTTACAGTGGCTCAGACTCCCTATCTGGGAGTGGGTTAGGTTCTTCAATCTACTA 299
QY 943 tcaaaagacacgccaagtgtgtggaattgtcaagaagcttcaactgctatgtatgc 1002
DB 300 TCAAAAGACACGCCCAAGTGTGGAATTTGTCAAGAGCTTTAACTGCTTACCTGATGC 359
QY 1003 tggggaggggtgtgttaaccattcgttaagcttggccggtgctggagataatgagacgctg 1062
DB 360 TGGGAGGCGGTGTACACCATTCGTAAAGCTTGCCTGGTGGACATATGACACGCTG 419
QY 1063 tggccctgatacggagacccctaatgagcttccatacaatgactacttgaatacttg 1122
DB 420 TGGCCCTGGATACGGAGATCCCTAATGAGCTTCATACATGACTCTTGAATTACTTGG 479
QY 1123 gaccgatttcaagctccacatcagtccttccaataatgactaaccagaaacgaatgagt 1182
DB 480 GACCAAGATTTCAGCTCCACATCAGTCTTCCAAATGACTTAACCGAGAACAGATGAGT 539
QY 1183 acctgagaagatcaaacgagcagctgtttgagaaccttgaatgctgcccagacagctg 1242
DB 540 ACCTGAGAGATCAAAACGACGACTGTTGAGAACCTTGAATGCTGCGCACGCACTG 599
QY 1243 gggctcaaacgagcagcttcttgaagagcagcattcccttgaagaaatgagaaagag 1302
DB 600 GGGTCAAAATGCAAGCGATTCCTGAGAGCGCCATCCCTAAGAGATGGCGATGAGAGG 659
QY 1303 aagaagaccctgacaagcagcatctgactgtctcctctgacaagaatctgcttgaag 1362
DB 660 AAGACAGACCTTGAACAAGCATCTGATCTGCTCTGTAACAACAAATTCCTGTGAGG 719
QY 1363 aagagttctccgattctgagaagagagagagggggcgcaagaactcttccaacttca 1422
DB 720 AAGAGTTCTCCATCTCTGAGAGAGGAGGAGGGGGCGCAAGAACTTTCACAACTTCA 779
QY 1423 aaaaagccaaagagagctcaaaaagagagatgaaaaagaaacccaagagagaaagag 1482
DB 780 AAAAAGCCAAAGAGATCAAAACAGAGATGAAAAAGAGAAAGCCAGAGAGAAAGAG 839
QY 1483 gaatcaccaagagagaaaccaaagagagagagcacaagcacaaggggttcaagagag 1542
DB 840 AAGTACCGAAGAGAGAG - AAACCAAGAGAGAGAGC - AGAAGCCAAAGGGGTCAAGAGAG 896
```

QY 1543 aggcacagtggcctgtaatgagac 1565
DB 897 GGTC--ARRTGGCTGAATGAGAC 917

RESULT 6
AL558916
LOCUS AL558916 847 bp mRNA linear EST 16-FEB-2001
DEFINITION AL558916 LTI_NFL008.TC2 Homo sapiens cDNA clone CS0DU007YP20 5
prime, mRNA sequence.
ACCESSION AL558916
VERSION AL558916.1 GI:12903904
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 847)
AUTHORS Li W.B., Gruber C., Jesse J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Genoscope - Centre National de Sequencage
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..847
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DU007YP20"
/clone_1ib="LTI_NFL008.TC2"
/sex="male"
/tissue_type="T cells from T cell leukemia"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com>"

BASE COUNT 217 a 195 c 224 g 207 t 4 others
ORIGIN

Query Match 52.1%; Score 839; DB 9; Length 847;
Best Local Similarity 99.1%; Pred. No. 3.9e-164;
Matches 839; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```
QY 101 ggcagcaagatgagcagacgagcagccggaagaagctgttactactacagag 160
DB 1 GCGGACCAAGATGCGGACACGAGGAGCCCGAGAGAAAGTCTGTACTACGACGG 60
QY 161 ggaagtgtgaattactattatgagcaagagccaacatgaagcctcacgaatccgat 220
DB 61 GGATGTTGGAATTTACTATTATGAGCAAGGCCAACCAATGAAGCCCTACCGATCGCAT 120
QY 221 gactcaaatgtgtcctaactatgtctcaccgaaaatggaatctatccgacctta 280
DB 121 GACTCAATATTTGCTGCTCAACTATGCTCTTACCAAAAATGGAATCTATGCGCCTTA 180
QY 281 caaagcaatgtcgaagagatgaccaagtaacacagcagatgactaatattcttcg 340
DB 181 CAAAGCAATGCTGAGAGATATGACCAAGTACCAACAGCATGACTATTAATTTCTGG 240
QY 341 ctccatcgcctcagataacatgctcgagtagcagcaagcagatgacagagatcacaagttgg 400
DB 241 CTCATCCCTCCAGATATGATGCGAGTACAGCAAGCATGACAGATGCAAGATTCACGTTGG 300
QY 401 tgaggacttccagatttgatggcctgtttgagttctgctgctcagtgctgaggttc 460
```


RESULT	9	954 bp	mrna	linear	EST 15-MAY-2001
LOCUS	BC754715				
DEFINITION	602714475F01 NIH_MGC_48 Homo sapiens CDNA clone IMAGE:4854547 5', mRNA sequence.				
ACCESSION	BC754715				
VERSION	BC754715.1	GI:14065368			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 954)				
TITLE	NIH-MGC http://mgc.ncl.nih.gov/.				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaabs-remail.nih.gov Tissue Procurement: Louis M. Staudt, M.D., Ph.D. CDNA Library Preparation: Ling Hong/Rubin Laboratory DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov plate: L1CML702 row: f column: 20 High quality sequence stop: 915.				
FEATURES					
source	Location/Qualifiers 1..954 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4854547" /clone_lib="NIH_MGC_48" /tissue_type="primary B-cells from tonsils (cell line)" /lab_host="DH10B (phage-resistant)" /note="Organ: B-cells; Vector: pOMB7; Site: 1: XhoI; Site: 2: EcoRI; CDNA made by oligo-dT priming. directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."				
BASE COUNT	249 a 214 c 265 g 226 t				
ORIGIN					
Query Match	49.08; Score 789.2; DB 10; Length 954;				
Best Local Similarity	94.0%; Pred. No. 8.8e-134;				
Matches 897; Conservative	0; Mismatches 48; Indels 9; Gaps 7;				
QY	96 gggagagcgcgaagatgcgcagacgacgagccagcgaggaagtctgttactactac 155				
Db	1 GGGGAGGGGACCAAGATGGCGACAGCGA -GGCACCCGGAGAAAGTCTGTACTACTAC 59				
QY	156 gacgaggaatgttgaanaattactatattatgacaagggccacccaatgaagcctcaacgaa 215				
Db	60 GACGGGGGATGTTGGAAATTAATATATGACAAAGGCCACCAATGAGCTCACCAGATC 119				
QY	216 cgcgatgataataatttgcctctcaa -ctatgctctcaccggaanaatgatattatcg 274				
Db	120 CGCATGACTCTAATATTCTCTCTCAAGCTAATGCTTCTACCGAAAAAGGAANTCTACG 179				
QY	275 cccccaacaagccaatgttgagagatgataccaagtaccacaagcgatgactcatataatt 334				
Db	180 CCCTCACAACCCCAATGCTTGAGAGATATACCAAGTACCAAGCCGATGACTTCATTAAAT 239				
QY	335 ctctgcgtcctcactcgtccagaataacatctgcgagctacgcaagcgagatgcagatca 394				
Db	240 CTTGGGCTCCATCGCTCCAGATTAACATCTGCGAGTACGACGAAGCATGTCAGAGATTCAA 299				
QY	395 cgttggtgaggaactgtccagatattccgatgagcctgttgatgtctgtcagttgtctactg 454				

Db 300 CATTGGTAGGACCTCCAGTAATTGCAATGCGCTGTTTGAAGTTCTGTCAATTGTCTACTACG 359
Qy 455 tggattctctggaagatgctgtgtaaaacttaataagaacgacgacgacatccgctgtaatg 514
Db 360 TGGTCTGTGGCAAGTGGCTGTGAACCTTAATTAAGACGACGACGACATCGCTGTGAATTG 419
Qy 515 ggcctggggcctgcgaacatgcaagaagtccgaagcatctggtcctctgttaacgtaaaga 574
Db 420 GGCTTGGGGGCGCTGCACCATATGCMAAAGATCCAGGACATCTGGCTCTCTGTAAAGTCAATGA 479
Qy 575 tatgctcttgagccatccctcggaaactctaaagatcatatcacaagaaggtgcgttaattgacat 634
Db 480 TATGCTCTTTGGCCATCTCGAAGCTCTTAAGATATACCAAGAGGTGCTGTACATTGACAT 539
Qy 635 tgaataaccatagttgacgcgctgtgaaagagagcctcttaacacgaagacccggtcatgac 694
Db 540 TGATATTACCACTAGTGAAGCGGTGGAAGAGCGCTTCTTACACACACGAGACCGGTCATGAC 599
Qy 655 tctgtcctcttcaataagatgagagtaacttcccaagaaactggggacctaaggatgatacgg 754
Db 600 TGTGTCCTTTCTTAAGATATGAGAGTACTCTCCAGGAACATGGGGACCTTAAGGGAATATCG 659
Qy 755 ggcctggaagaacaagatattatgctgttaaacccgcctccgaagacggatgatacga-tgacg 813
Db 660 GCGTGGCAAGGCAAGTATATAGCTGTATACATCCCGCTCCGACAGAGGATTTGATTGACG 719
Qy 814 agtccatagagccatttccaagccggtcatgtccaaa--gtaatggagatgttccag-- 869
Db 720 AGTCTATGAGAGCCATTTTCAAGCGGACATGTCCAAAGATTAAGAGATGTTCACAGAC 779
Qy 870 cctagtcggtgtggtcttaacagtgtg-gcctgaacccctcatcttgggagtgatggtg 928
Db 780 CTAGATGGGGTGTCTTAACAGAGGTGAGTCAAACTCTTATCTGGGGATCGGGAAGGTGG 839
Qy 929 ctcaactatcatcaagaagacacgccaaggtgtgtgaattgtccaagagcttaacct 988
Db 840 TTCACCTCTAATCAAGGAAGAACCCCAAGGTGTGTGCAATTGTCCMAAGC-TTAACTT 898
Qy 989 gactatgctgtagctctggagagcggtgtgttaacacatctgttaacgtgtcccggtg 1042
Db 899 GCGTATGCTGATGCTGGGAAGGGGTGTAAACCATTTTCGTAACTGTACGCTG 952

RESULT 10
68756668
LOCUS 60271552/F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4855663 5',
DEFINITION mRNA sequence.
ACCESSION BG756668
VERSION BG756668
KEYWORDS BG756668.1 GI:14067321
SOURCE EST.
ORGANISM human.
REFERENCE Homo sapiens
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL NIH-MGC <http://mhc.nci.nih.gov/>.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps@remail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
<http://image.llnl.gov>
 plate: LCM1705 row: e column: 08
 High quality sequence stop: 818.
 Location/Qualifiers
 1. 910

FEATURES
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4855663"
/clone_1lb="NIH_MGC_48"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: B-cells; Vector: pOT7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT      231 a      201 c      253 g      224 t      1 others
ORIGIN
Query Match      48.4%; Score 779.4; DB 10; Length 910;
Best Local Similarity 95.7%; Pred. No. 9.4e-152;
Matches 833; Conservative 0; Mismatches 32; Indels 5; Gaps 3;
439 gtcagtgctcactggtgtgttcctgtggaagtgctgtaacttaagcagcaagc 498
|||||
2 gtcagtgctcactggtgtgttcctgtggaagtgctgtaacttaagcagcaagc 61
|||||
499 acatcgccgtgaattgggtgtgggctgcacatcgaagaagtcgagcctcgctc 558
|||||
62 acatcgccgtgaattgggtgtgggctgcacatcgaagaagtcgagcctcgctc 121
|||||
559 tctgttaagtcgaatgatgatcgtcttgccatcctggaactgtctaaagatcaccagaag 618
|||||
122 tctgttaagtcgaatgatgatcgtcttgccatcctggaactgtctaaagatcaccagaag 181
|||||
619 tctgttaagtcgaatgatgatcgtcttgccatcctggaactgtctaaagatcaccagaag 678
|||||
182 tctgttaagtcgaatgatgatcgtcttgccatcctggaactgtctaaagatcaccagaag 241
|||||
679 cgagccgggtcactgactgtccttcataagatagaaagatcctccaggaactgg 738
|||||
242 cgagccgggtcactgactgtccttcataagatagaaagatcctccaggaactgg 301
|||||
739 acttaaggatatacggggtgtgcaagaagaatgataatgtctgttaactaccgctcgag 798
|||||
302 accttaaggatatacggggtgtgcaagaagaatgataatgtctgttaactaccgctcgag 361
|||||
799 acggagatgatagagctcctatagagccatttaagccggtcattcgtccaaagatag 858
|||||
362 acggagatgatagagctcctatagagccatttaagccggtcattcgtccaaagatag 421
|||||
859 agatgttcacagcctagtcgtgtgtcttacaagtgctcagaactccctatctggagatc 918
|||||
422 agatgttcacagcctagtcgtgtgtcttacaagtgctcagaactccctatctggagatc 481
|||||
919 ggttaagtgcttcaatctatctatacaagaagcagcgaagtgctggaattgttcaaga 978
|||||
482 ggttaagtgcttcaatctatctatacaagaagcagcgaagtgctggaattgttcaaga 541
|||||
979 gcttaacatcgtcctatgctgagtcgtggagggcggtgttaacacatctgtaagctggccc 1038
|||||
542 gcttaacatcgtcctatgctgagtcgtggagggcggtgttaacacatctgtaagctggccc 601
|||||
1039 ggtgtcgtgacatacgagacagctgtggccttgatacggagatccctaaatgaagcttccat 1098
|||||
602 ggtgtcgtgacatacgagacagctgtggccttgatacggagatccctaaatgaagcttccat 661
|||||
1099 acatatacacttgaatctctttggaccagattcaagctccacatcagctccttccaata 1158
|||||
662 acatatacacttgaatctctttggaccagattcaagctccacatcagctccttccaata 721
|||||
1159 tgactaccagaaca--cgatagtagacttggaagaatcaaacagcgactgtttgaaga 1216
|||||

```

```

Db 722 TGACTAACGAGAACAAAGAAATGACTGTGAGAGATCAACAG-GACTGTTTGAGAA 780
Oy 1217 cctagaatgctgcgcagcagcactggtgggtccaaacagcag--cgattcctgagcagcc 1274
|||||
Db 781 CCTTAAATAATGCTTGGCGACGACACTGGGCTCCAAATGCAAGGGGATTCCTTAAGACGCC 840
Oy 1275 atccctgagagagatgctgcatgagagcga 1304
|||||
Db 841 ATTCTGAGAGAGATGGGATGAGGAGCA 870
|||||

RESULT 11
BM477303
LOCUS      BM477303      982 bp      mRNA      linear      EST 05-FEB-2002
DEFINITION AGENCOURT_6485091 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5554070
5', mRNA sequence.
ACCESSION  BM477303
VERSION     BM477303.1 GI:18526345
KEYWORDS   EST.
SOURCE      human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 982)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            JOURNAL
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgrabs-remail.nih.gov
            Tissue Procurement: Lou Staudt
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/ILNI at:
            http://image.llnl.gov
            Plate: LLM12272 row: a column: 15
            High quality sequence stop: 677.
            Location/Qualifiers
                source          1..982
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5554070"
                /clone_1lb="NIH_MGC_85"
                /tissue_type="Lymphoma, cell line"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: Lymph; Vector: pCMV-Sport6; Site_1: NotI;
                Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                Average insert size 1.867 kb. Library enriched for
                full-length clones and constructed by Life Technologies.
                Note: this is a NIH_MGC Library."
BASE COUNT      251 a      225 c      269 g      235 t      2 others
ORIGIN
Query Match      48.3%; Score 778; DB 10; Length 982;
Best Local Similarity 96.7%; Pred. No. 1.9e-151;
Matches 847; Conservative 0; Mismatches 22; Indels 7; Gaps 5;
97 gggagcggagcgaagatggcgagcagcagcggcggggaagctttactactag 156
|||||
Db 19 gggagcggagcgaagatggcgagcagcagcggcggggaagctttactactag 78
|||||
Oy 157 acgggagatgttggaaattactatatacgacaagcaccacaaatgaagcctcaccagatcc 216
|||||
Db 79 ACGGGATGTTGAATTTACTATTATGACAAAGCCACCAATGAAGCCTCAACCGAATCC 138
|||||
Oy 217 gcatgactataatgtctgctcactaactatgctctcaccgaaatgtgaatctatcgcc 276
|||||
Db 139 GCATGACTCATATATTGCTGCTCAACTATAGTCTCTCAACGAAATGAATCTATCGCC 198
|||||
Oy 277 ctacaaagcgaatgtctggagatgacaaagcagcgatgctactaatctc 336
|||||

```

DB 199 CTCACAAAGCCAAATGCTGAGAGAGATGACCAAGTACACAGCATGACTTAATTAATTC 258
 QY 337 tgcgcctccatccgtccagataacatgctcgagatcgagcaagcagatccagatccaa 396
 DB 259 TCGGCTCCATTCGCTCCAGTAATACATTCGAGTACAGCAAGCATGATCCAGATTCAC 318
 QY 397 ttggtgaggaagctgctcagatattcgatgctgttggatgctgctgctgctgctgct 456
 DB 319 TTGGTAGAGACTGTCAGATTCGATGAGGCTGTTGAGTTCCTGCTGATTCCTACGCTG 378
 QY 457 gtctctgtgagcagctgtgtggaacttaataagcagcagcagcagcagcagcagcagc 516
 DB 379 GTTCTGTGCAAGTGTGTAAGAACTTAATTAAGCAGCAGCAGCAGCAGCAGCAGCAGC 438
 QY 517 ctggagggcctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 576
 DB 439 CTGGGGGCTGCACCATGCAAGAACTCCGAGCATCTGCTTCTGTTACGTCAATGATNA 498
 QY 577 tgcctgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 636
 DB 499 TCGTCTTGGCCATCCCTGGAACCTGCTAAAGTATACACAGAGGCTGCTACATGACAT 558
 QY 637 atattccatctgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 696
 DB 559 ATATTTCACCATGCTGAGCGGTGGAAGAGCCCTTCTACACACAGCAGCAGCAGCAGC 618
 QY 697 tgcctctataagatgagagatctccaggaactggagagcagcagcagcagcagcagc 756
 DB 619 TGTCTTTATATAAGTATGAGAGTACTTCCAGGAACCTGGGACCTACCGGATATCGGG 678
 QY 757 ctggcagaagcagatattatgctgttaactaccgctcagcagcagcagcagcagcagc 816
 DB 679 CTGGCAAAAGCAAGTATTTGCTGTAACTACCCGCTCCGAGACCGGATGATGAGCAG 738
 QY 817 cctatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 874
 DB 739 CCTATGAGCCATTTTCAAGCCGCTATCTCCAAAGTATGAGATGTTCCAGCCTAG 798
 QY 875 tgcggt-aggctctacagt-gtgcctcagcagcagcagcagcagcagcagcagcagc 929
 DB 799 TCGGAGGCTGTACAGTGTGCTGCTGAGACTCCCTATCTGAGNGATCCGTTAGGCTTG 858
 QY 930 ttcaatctatctatacaagagcagcagcagcagcagcagcagcagcagcagcagc 965
 DB 859 TTCAATCTAATCTATTCAGGAACCCCAAGNMTG 894

RESULT 12
 BI772704 778 bp mRNA linear EST 25-SEP-2001
 LOCUS 603053373P1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5202739 5',
 DEFINITION mRNA sequence.
 ACCESSION BI772704
 VERSION BI772704.1 GI:15764282
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bcrfemail.nih.gov
 Tissue Procurement Life Technologies, Inc.
 cDNA Library Preparation Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM11508 row: b column: 20

High quality sequence start: 6
 High quality sequence stop: 762.
 Location/Qualifiers
 1. 778
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5202739"
 /clone.lib="NIH_MGC_122"
 /lab_host="DH10B"
 /note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
 Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
 anonymous pool of 24 week female lung, 16 week female
 spleen, and 20-22 week male spleens. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.4 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 026. Note:
 this is a NIH-MGC Library."
 BASE COUNT 201 a 177 c 207 g 193 t
 ORIGIN
 Query Match 45.8%; Score 737.4; DB 10; Length 778;
 Best Local Similarity 99.0%; Pred. No. 4.8e-143;
 Matches 763; Conservative 0; Mismatches 6; Indels 2; Gaps 2;
 QY 136 ggaagctgttactactcagcagcagcagcagcagcagcagcagcagcagcagcagc 195
 DB 8 GGAAGCTGTGTAATCTACGAGGGGAGTGGAAATTAATTAATTAATTAATTAATTAAT 67
 QY 196 caatgagcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 254
 DB 68 CAATGAGCCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 127
 QY 255 cgaagaaatgaaatctatcagcagcagcagcagcagcagcagcagcagcagcagcagc 314
 DB 128 CGAAATATGAAATCTATGCTTCACAAAGCTAAAGCTGAGAGATGACCAAGTACAC 187
 QY 315 agcagatgaactaactaactcagcagcagcagcagcagcagcagcagcagcagcagc 374
 DB 188 AGCGATGACTATTAATTTCTTGGCTTCATCCGATATGATGATGATGATGATGATG 247
 QY 375 aagcagatgcaagatcaacgcttggatgagcagcagcagcagcagcagcagcagcagc 434
 DB 248 AAGCAGATCAGAGATTAACGTTGGTGAGACTGTCAGATATTCATGAGCGCTTGG 307
 QY 435 ttctgtcagttgtctactcagcagcagcagcagcagcagcagcagcagcagcagcagc 494
 DB 308 TTCTGTCAAGTTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 367
 QY 495 acgagatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 554
 DB 368 ACGGATCAGCTGTGATTTGGCTGGGCTGACCATGCAAAAGATGCCAGCATCT 427
 QY 555 ggcctctgttactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 614
 DB 428 GGGTCTGTAGCTCAATGATATCGCTTGGCATCTGCAATCTGCAATCTGCAATCTG 487
 QY 615 aggtgtcttactcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 674
 DB 488 AGGGTCTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 547
 QY 675 accaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 734
 DB 548 ACCAGCAGACGGGCTATGACTGTGCTTCTTCAATGATGAGAGATCTCCAGAGACT 607
 QY 735 gggagcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 794
 DB 608 GGGAGCTACGGGATATCGGGCTGCAAGCAAGTATTAATTAATTAATTAATTAATTA 667
 QY 795 cgagacggatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 854

Db 668 CGAGACGGGATTGATGACGAGTCTATAGAGCCA-TTTCAMGCCGGTATGCCAAAGTA 726

Qy 855 atggaatgttcacagccctagctgagtggtctctacagtgctgctcagactcc 905
|||||
Db 727 ATGGAAGTGTTCACACCTTAGTCCGGGGCTTTACAGTGTGCTCAGACTCC 777

RESULT 13
BM452740 1007 bp mRNA linear EST 05-FEB-2002
LOCUS BM452740
DEFINITION AGENCOURT_6401305 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5498480
5', mRNA sequence.
ACCESSION BM452740
VERSION BM452740.1 GI:18501780
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1007)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Lou Staudt
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LLM12130 row: e column: 09
High quality sequence stop: 558.
Location/Qualifiers
1. 1007
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5498480"
/clone_lib="NIH_MGC_85"
/tissue_type="lymphoma, cell line"
/note="Organ: lymph, Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dr primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."
BASE COUNT 256 a 259 c 257 g 235 t
ORIGIN

Query Match 45.7%; Score 736.8; DB 10; Length 1007;
Best Local Similarity 97.5%; Pred. No. 6.8e-143;
Matches 770; Conservative 0; Mismatches 17; Indels 3; Gaps 2;

Qy 405 gactgtcagatattcgatgctgtttagtctgtcagtgctcactcgtggtctctg 464
Db 1 GACTGTCCAGATTGATGAGGCGCTTTGAGTTCTCTCAGTTGCTTACCTGCTGTTCTG 60

Qy 465 gcaatgtctgtaaaactaataagacagacagacatccgtgaattggctggggc 524
Db 61 GCAATGTCTGTGAACCTTAATAAGACAGACGACATGCTGTGAATTTGGCTGGGGC 120

Qy 525 ctgacacatgcaaaagatccagagcatctgctctcttcaagcaatcgtctctg 584
Db 121 CTGACACATGCAAAAGATCCGAGGACATCTGCTCTCTTACGTCAATGATATCGCTCT 180

Qy 585 gccatcctggaactgtaagatatacagagaggtgctgtatattgacattgataac 644
Db 181 GCCATCTCTGGAAGTCTTAAGATATCACAGAGGCTGTGTATATGACATTAATTCAC 240

Qy 645 catgtgacggtggtggaagagcctctacaccagagccggttatgactgtgtccctt 704
|||||

Db 241 CATGTGACGGCGTGGAAAGAGGCGCTTTACACACGAGACCGGGGTATGATGTGCTTT 300

Qy 705 cataagatggaagatctctccaggaactgaggacctaagagataccgggtctggcaaa 764
|||||
Db 301 CATTAAGTATGAGAGTACTCTCCAGAGACTGGGAGACCTACGAGATATGGGGCTGGCAA 360

Qy 765 gacaagattatgtctgtaactaccgcgtccgagacgggattgataagagtcctatgag 824
|||||
Db 361 GGCAAGTATTTAGTGTATTAACATACCCGCTCCGAGACGGGATGATGAGAGCTCATGAG 420

Qy 825 gccatttccagccggtatgtcccaagtaagatgagatgttccagccttagtcgtgtc 884
|||||
Db 421 GCCATTTTCAAGCCGATCATGTCCAAAGTAATGAGATGTTCCACCTAGTGTGGTGGTC 480

Qy 885 ttacagtggtgtcagactccctatctgggagtcggttagtggttcaatcattatc 944
|||||
Db 481 TTACAGTGTGGCTCAGACTCCCTATCTGGGAGATCGGTATGTTGCTTCATATTAACAT 540

Qy 945 aaagacacgccaagtgtgtggaattgtcaagagccttaacctatgctgtatgctg 1004
|||||
Db 541 AAAGACACGCGCAAGTGTGTGGAATTTGCAAGACCTTTAACCTCTATGCTGATGCTG 600

Qy 1005 ggaagcgtgtgttaccacatctgtacagtggtccggtgtgtgacatatgagacgtgtg 1064
|||||
Db 601 GGAAGCGGTGTGTACACCACTTGTACAGTTGCTGACCTGCTGACATATGACACTGTG 660

Qy 1065 gccctgtatcagagatccctaatgagcttccatacaatgactattgaaactttgga 1124
|||||
Db 661 GCCCTGATGATGAGAGATCCCTAATGAGCTCCATFACAAATGACTACTTTGAATACCTTTGGA 720

Qy 1125 ccagatttcaagctccaatcag-tccttccaatatgactaacag--aacagaaatgag 1181
|||||
Db 721 CCAGATTTCAACCTCCACATCAGACCTCCCAATATGACTACCCGGAACCCCAATGAG 780

Qy 1182 taccctggaga 1191
|||||
Db 781 TACCTGGGA 790

RESULT 14
BG768807 803 bp mRNA linear EST 15-MAY-2001
LOCUS BG768807
DEFINITION 602743069P1 NIH_MGC_49 Homo sapiens cDNA clone IMAGE:4872689 5',
mRNA sequence.
ACCESSION BG768807
VERSION BG768807.1 GI:14079460
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 803)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LLM1749 row: j column: 18
High quality sequence stop: 801.
Location/Qualifiers
1. 803
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4872689"
/clone_lib="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"


```

OY 517 ctgggggctgacatgcaagaagtcgaggcatctgctctgttcaatgata 576
    |||
DB 441 ctgggggctgacatgcaagaagtcgaggcatctgctctgttcaatgata 500
    |||
OY 577 tctcttgccatctggaactgctaaagtatcacagaggtgctgtacatlgacatlg 636
    |||
DB 501 tctcttgccatctggaactgctaaagtatcacagaggtgctgtacatlgacatlg 560
    |||
OY 637 atattcacatgltgacggcgtlgaagagcctctacacacagacgggtcatgactg 696
    |||
DB 561 atattcacatgltgacggcgtlgaagagcctctacacacagacgggtcatgactg 620
    |||
OY 697 tctcttgccatgagatgagagtgacttccaggaacttgaggacttaacgggataccgggg 756
    |||
DB 621 tctcttgccatgagatgagagtgacttccaggaacttgaggacttaacgggataccgggg 680
    |||
OY 757 ctggcaagacaagatatactgtltaactacacggctccgagacggatlgatgacagt 816
    |||
DB 681 ctggcaagacaagatatactgtltaactacacggctccgagacggatlgatgacagt 740
    |||
OY 817 cc--tatgaggccatttcaagcgggtcatgtccaaglaatgagatgttcc 867
    |||
    741 ccttatgaagcatttttcaaccggcctgagcctcaaaagtaagagaggttcc 793
    |||

```

Search completed: June 8, 2002, 11:39:48
 Job time: 7656 sec



```

/protein_id="AAC50475.1"
/db_xref="GI:1277084"
/translation="MAOTGTRRKYCYIIDGVDGNYVYGGDHMKPHRIIRMTNLLN
YGLIRMEYIRPRHNAEMETKYSHSDYDKLPRLSPDPMKSEKMGQSEYNDGECPV
PGLEPERCOLSTGGSYASAVKLNKQOTDIYVNRMGVLSFKYGEYEPFGTDLDIDGK
LELTKHQREVLITIDIDIHGGDVEEAFTYTDNRMTVSFKYGEYEPFGTDLDIDGK
GKYAYVPLRDGIDSESEALFPRMYSVMEMCFQSAVLOGSDLSGDRIGENL
LTIGHAKCYELPFGKSFNLPMILMGGGGYTLIRNARCPYTLVAALDTEINELPLNDYF
EYEGPDRFKLSPSSNNTNONTNLEYELIKORLPENLRMLPHAGYOMALIPDAIDEE
SDEDEDDBDKRSTLSPSSDKRTIACGEEPSDESEEGCGKAKNSNFKAKRVTDEDEK
KDPKEKVEYTEEKTEEKPEAKGYAEVYKLA"

```

Query Match	99.18;	Score 1596.6;	DB 9;	Length 1611;
Best Local Similarity	99.48;	Pred. No. 0;		
Matches 1602; Conservative	0;	Mismatches 9;	Indels 0;	Gaps 0;

QY	1	atgtctcggggtctctgcgcgctggtgtctgtctgtctccactcggcactctgaagaca	60
Db	1	ATGTCTCGGGGTCTCTGCTGCCCTGGTGTGCTGTCTCCACTCGGATCTCTGGAACACA	60
QY	61	gctctgagcgrctctgttcacctcgggtgtgagaccacgcgggggggagagaaatgtgcgaga	120
Db	61	GCTTGACCGTCTCTGTCACTCGGGGTGAGACACGCGGGGAGGAGCAAAATGACGAGA	120
QY	121	cgcagagccacccggaggaagtcgtcttactactacagacgggagatgtttggaatattact	180
Db	121	CGCAGGCGACCCGGAGGAAGTGTGTACTACTACAGCGGGGATGTTGGAATTAATTA	180
QY	181	atggaaacagggccaccacaatgaagcctccacccgaatccgatgtactcataatttctgtca	240
Db	181	ATGGACAAGGCCACCCAAATGAACCTTCACGAATCCGCAATCCGATCAATTAATTCCTGCTCA	240
QY	241	acatagttctctaccgaaataatgaaatctatctatgccttaccgaagccaatgctcggagaga	300
Db	241	ACATAGTCTCTACCGGAAAAATGAAATCTATGCCCCCTCACAAAGCCAAATGCTGAGAGA	300
QY	301	tgaaccaagtaccacagcgaatgactacaatlaattcttgcctccatccgcgtccagataaca	360
Db	301	TGACCAAGTACCAACAGGATGACTACATTAATTTCTTGCGCTCCATCCGTCAGATATACA	360
QY	361	tgtcgcgaatcacgaacagacagatgcagagatcaacagtttgttgaagactgtccaglatcg	420
Db	361	TGTCGGAGTACACCAAGCAGATGCAGAGATTCAAAGTTGGTGAAGACTGTCCAGTATTCG	420
QY	421	atggccgttttgagttctgtcgaattgtctactgtgtgtttcttgygcaagtgctgtgaac	480
Db	421	ATGGCCGTTTGAGTTCTGTCAATGTGTCAATCTGTTGTTGGCAAGTGCTGTGAAC	480
QY	481	ttataaagcagcagacgagacatcgccgttgatctgggcttggggccctgcacatgtacaaga	540
Db	481	TTATATAAGCAGCAGACGAGACATCGCTGTGAATTTGGGCTGGGGGCTTCACATGTCAAAGA	540
QY	541	agtcgcgaagcactgtgctctgttcaacgtcaagatgatactcgtcttgygcacatccctgaa	600
Db	541	AGTCCGAGCATGTGGCTTCTGTACGTCGAATGATATGCTTGGGCATCTCTGGAATGTC	600
QY	601	taaaagtatacccaaggggtgtctgtacattgtacattgatataccaatgtgtgacggcggtg	660
Db	601	TAAAGTATCCACGAGGGTGTCTGTACATTTGATTTATTCACAAATGTCAGACGGGCTGG	660
QY	661	aagagagccttcaacaacagagacccgggtctatgactgtgtccttcaaatatgtgagagt	720
Db	661	AAGAGAGCCTTCTACACACGAGGACCGGGTCTATGACTGTGTCTTTTCAATAATAGAGAGT	720
QY	721	acttcccaaggaacttgggaacttaacgggataccggggcttgcacaaagaacgaatattatgtc	780
Db	721	ACTTCCCAAGAACTGGGAGACTTAAGGATATCGGGGCTTGCAAAAGCAATATTATAGCTG	780
QY	781	ttactacccgcctcccgagagcggaattgatagtacagatcctatgtggccatttcaagccgg	840
Db	781	TTACTACCCGCCTCCCGAGAGCGGAATTGATAGTACAGATCCTATGTGGCATTTCAGGCCGG	840

Db	781	TTAAGTACCCGGCTCCGAGAGCGGGATTGATGACAGTGTCTTAAGAGCCATTTCAAAGCCG	840
Qy	841	lcatgtccaaagatgatgagatggttccagccatgtgcgtgtgttcaagtggtctcag	900
Db	841	TCATGTCCAAAGATGATGAGATGTTCAGGCTAGTGGGGTGTCTTACAGTGTGGCTAG	900
Qy	901	actccctatctggggtatcggttaggtgtgtcctaattcatctatcataaagacagccaaat	960
Db	901	ACTCCCTATCTGGGGATCGGTTAGTGTCTCAATCTTAACATACAAAGACACCCCAAGT	960
Qy	961	gtgtgagatttgtcaaaagctttaactctgctatgctgctgctgtggaggcggtgttaca	1020
Db	961	GTGTGAGATTGTGTCAAAAGCTTTAACTGCTTATGCTGATGCTGTGGAGCGGTGTTCACA	1020
Qy	1021	ccattgtcaagctgtgcccggtgtgtgcacatactgaaagacgcgtgtgacctgtatatcgaga	1080
Db	1021	CCATTGCTTAACGTTGGCCCGGTGCTGTGACATATGAGACAGCTGTGGCCCTGGATACGGAGA	1080
Qy	1081	tcacctatgagcttccatacactgtactacttgtgatacttggacattggacagattcaagctcc	1140
Db	1081	TCCTATATGAGCTTCCTACATGACATGACTTGAATATCTTGGACACAGATTTCAAAGCTCC	1140
Qy	1141	acatcagctcctccacatactgactaaccaagaaacagcaatgagctgagagaagatccaac	1200
Db	1141	ACATCAAGCTCTTCCAAATATGACATTAACCAAGAACAGCAATGATGACTGTGAGAGATCAAC	1200
Qy	1201	agcgacatgtttggaagacctagaatgctgtgcgcagcgcaacttgggtgtccaaagcgaagcgga	1260
Db	1201	AGCGACATGTTTGAAAMACCTTAGAATGTGTGCGCACGCACTCGGGGTCCAAATGAGAGGGGA	1260
Qy	1261	ttctctggaagcgccatccctcctgaagagagtgtcgcatggaagacgaagaccccttcaaacg	1320
Db	1261	TTCTCTGAGAGCCCATCTCCCTGAGGAGTGGCGCATGAGAGACGAAGACCCCTGTCAAGC	1320
Qy	1321	gcatctgcatctgctcctcttgcacaaagaaattgtcctgtgagaaagagttcttcgattctg	1380
Db	1321	GCATCTGCATCTGCTCTCTGTGACAAAGAAATGTGCTGTGAGAGAAAGATGTTCCCGATTGTG	1380
Qy	1381	aagaagaggaagagggggccgcagaagacttccaattcaaaaaagccaagagatca	1440
Db	1381	AAGAGGAGGAGAGGGGGGCGCGCAAGAACTTCCAACTTCAAAAAAGCCCAAGAGATCA	1440
Qy	1441	aaacagagagatgaaaaaagagaagaaccacagagagagaagaagaaagatacccgaaagagaga	1500
Db	1441	AAACAGAGGATGAAAAAGAGAAACCACAGGGAGGAAGAAAGAAAGTCAACCCAAAGAGGAGA	1500
Qy	1501	aaaccaagagagaagaacccagaagccaaagagggtccaaagagaagagccaagtgtgacctbaa	1560
Db	1501	AAACCAAGGAGGAACCCAGAACCCAAAGGGGTCAAGAGAGAGATCAAGTTGGCCTTGA	1560
Qy	1561	tggagacctccagctcgtgcttccgtgtgagttccctcaagctttcttccccc	1611
Db	1561	TGGAGCCTCTCAGCTCTGGGCTCTCTGCTGAGTCCCTCAACGTTTCTTTCCCC	1611

RESULT	3
LOCUS	BC000301
DEFINITION	Homo sapiens, histone deacetylase 1, clone MGC:8378 IMAGE:2820260,
ACCESSION	BC000301
VERSION	BC000301
KEYWORDS	MGC.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2093)
REFERENCE	Strausberg, R. Direct Submission. Submitted (15-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
TITLE	
JOURNAL	

C



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2002, 09:35:07 ; Search time 3028.56 Seconds
(without alignments)
11131.578 Million cell updates/sec

Title: US-09-817-538-2

Perfect score: 1611
Sequence: 1 atgtctggggctctctgccg.....tccctcagcttcttccccc 1611

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 segs, 10463268293 residues

Minimum number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vi: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htgo_inv: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1596.6	99.1	1611	6	AX053101
2	1596.6	99.1	1611	9	HS050079
3	1499.6	93.1	2093	9	BC000301
4	1498	93.0	2091	9	D50405
5	1498	93.0	2111	6	AR012007
6	1498	93.0	2111	6	162388
7	1498	93.0	2111	23	E11455
8	1240.2	77.0	1977	10	NH0145
9	1236.4	76.7	1999	10	MMU80780
10	1110.4	68.9	81925	2	AP000450
11	1110.4	68.2	146841	2	AC020549
12	1098.4	68.2	187852	2	AC068041
13	1054	65.4	252957	2	AC105705
14	997.4	61.9	1646	5	AF044169
15	997.4	61.9	1678	5	AF039751
16	992.6	59.6	1646	5	AF043328
17	953	59.2	1487	5	AF020658
18	943.4	58.6	2305	5	XLAB21
19	873.6	54.2	3701	5	CHRTCKL
20	873.6	54.2	4073	5	GGTK1
21	821.2	51.0	1878	5	AF039752
22	818	50.8	1997	6	AX05742
23	818	50.8	1997	10	MMU31758
24	812.8	50.8	101767	9	AL389885
25	811.2	50.4	140242	2	AC073038
26	805.2	50.0	1985	6	AX053103
27	805.2	50.0	1985	6	AX053103
28	752.2	46.7	121601	9	AL592406
29	752.2	46.7	154386	2	AC016314
30	752.2	46.7	159286	2	AC025710
31	734.4	45.6	2180	3	AF032919
32	729	45.3	2157	3	AY058487
33	717.8	44.6	2172	3	DMH1SEAC
34	712.8	44.2	2073	3	AF026949
35	655.8	40.7	106840	2	AC094420
36	654.6	40.6	70949	2	AC095666
37	642.2	39.9	2681	3	AF086751
38	640.6	39.8	51246	2	AC014151
39	640.6	39.8	156525	3	AC010665
40	640.6	39.8	301769	3	AE003482
41	611.8	38.0	768	10	AF321129
42	583	36.2	45754	9	AL139118
43	551.6	34.2	158526	2	AC094458
44	544.2	33.8	106731	2	AC097622
45	532	33.0	109924	2	AC105353

ALIGNMENTS

RESULT 1
AX053101
LOCUS AX053101
DEFINITION Sequence 25 from Patent WO0071703.
ACCESSION AX053101
VERSION AX053101.1 GI:12227157
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1611)
AUTHORS Macleod, A.R., Li, Z. and Besterman, J.M.
TITLE Inhibition of histone deacetylase
JOURNAL Patent: WO 0071703-A 25 30-NOV-2000;
Methylene, Inc. (CA)
FEATURES
Location/Qualifiers
source
1..1611
/organism="Homo sapiens"
/db_xref="taxon:9606"

linear PAT 12-JAN-2001

BASE COUNT 428 a 385 c 440 g 358 t
ORIGIN

Query Match 93.0%, Score 1498, DB 9, Length 2091.

		Best LocalSimilarity	99.3%;	Pred. No. 0;	Mismatches	10;	Indels	0;	Gaps	0;
		Matches	1504;	Conservative	0;					
Oy	97	gggagcgagcaagaatgagcgacagcgagcgagccagaggaagtctgtactactacg	156							
Db	50	gggagggcagcagcaagatggggcgacagcgagggcagccgggagaaatcttctactacg	109							
Oy	157	acgggagtggttgaatctactatitgtgacaagggccacccaatgaagctccacgaatcc	216							
Db	110	acggggagtttggaattacttcttatttgacaaagccaccatgaacctcaccgaatcc	169							
Oy	217	gcatcactcataatgtgtcgtctcaactabgtctctacccaaaatlynaatctalcgcc	276							
Db	170	gcattacatcatatttggctgctcaactatgtctctacccgaataatgaaatctatcgcc	229							
Oy	277	ctcaaaagccaaatgtctgaagagatgacccaagtaccacaagcgatgactacataattct	336							
Db	230	ctcacaaaagccaatcttgagagatgaccaaatgaccaaatgaccatgacatcaattaaattct	289							
Oy	337	tgagctccatccgctccagataaacaatgctcgagatagacagaagaagatgacagatctcaacg	396							
Db	290	tgccgtccatccgctccagatatacagtgtagctacgcaagcaagatgacagatctcaacg	349							
Oy	397	tttgtgaagactgtctccagatctcgatctgagcgttttgagttctgtcaagtgtctactgtgt	456							
Db	350	tttgagagagactgtctcagatttcgatttcgatttcgatttcgatttcgatttcgatttcg	409							
Oy	457	gttcgttggaagtgctgttgaacttaataagcagacagacagacatcgccgtgaattgg	516							
Db	410	gttcgttggaagtgctgttgaacttaataagcagacagacagacatcgccgtgaattgg	469							
Oy	517	ctgggggcccctgcacatgcgaagaaggtccgagacatctgtctctgttaccgtcaatgata	576							
Db	470	ctgggggcccctgcacatgcgaagaaggtccgagacatctgtctctgttaccgtcaatgata	529							
Oy	577	tcgtcttgagccatctcttggaactgcttcaaaagtatcaccaagaggtgtctgtacatgtgaatg	636							
Db	530	tcgtcttgagccatctcttggaactgcttcaaaagtatcaccaagaggtgtctgtacatgtgaatg	589							
Oy	637	atatcaccaatggtgacggtcgttgaagaagcctcttacaaccaagacacgggtcatgactgt	696							
Db	590	atatcaccaatggtgacggtcgttgaagaagcctcttacaaccaagacacgggtcatgactgt	649							
Oy	697	gttcccttcataagatgtagagtgacttcccagagacttgggaactacaggttacccgggt	756							
Db	650	gttcccttcataagatgtagagtgacttcccagagacttgggaactacaggttacccgggt	709							
Oy	757	ctggacaagaacaaagtatatgtctgttacaacccgctccgagaacggatctgtatgaacgt	816							
Db	710	ctggacaagaacaaagtatatgtctgttacaacccgctccgagaacggatctgtatgaacgt	769							
Oy	817	ccctatgagggcaatttccaagccggtcaatgctccaaagtaatgtagatgttccaagctagt	876							
Db	770	ccctatgagggcaatttccaagccggtcaatgctccaaagtaatgtagatgttccaagctagt	829							
Oy	877	cggttggtcttaagaagtgtgtcagagctccctctctggagatcggttgaagtgttcttcaatc	936							
Db	830	cggttggtctttagagtggtgtcagagctccctctctggagatcggttgaagtgttcttcaatc	889							
Oy	937	tatctacaagaagacacgcccgaagtgtgttgaatttgttcaagaagctttaaaccgtctatgc	996							
Db	890	taactatcaataaagacacgcccgaagtgtgttgaatttgttcaagaagctttaaaccgtctatgc	949							
Oy	997	tgtatgtctggagagcggtgtgttatacacaattctgttaacggttgccgggtgctgtagaatatgaga	1056							
Db	950	tgtatgtctggagagcggtgtgttatacacaattctgttaacggttgccgggtgctgtagaatatgaga	1009							
Oy	1057	cagctgtgtgcccctgtgatacagatccctcaatgagcttccatacaatagtactcttgaat	1116							
Db	1010	cagctgtgtgcccctgtgatacagatccctcaatgagcttccatacaatagtactcttgaat	1069							
Oy	1117	actttggaccagatttcaagctccaatcagcttcttccaatatgtactaaacgaacaga	1176							

Db 1070 ACTTTGGACAGATTTCAGCTCCACATCATGCTCTTCCAAATATGACTAAACAGAACGCA 1129
Qy 1177 atgagtaacttgagaagaatcaaacagcgaactgtttgagaaccttaagaacttcgcgaacg 1236
Db 1130 ATGAGTACTGGAGAGATCAAAACAGCACTGTTTGAAGACCTTAAGAAATGCTGCCCCAGC 1189
Qy 1237 caacttgggttccaacacgcagcagatctccttgaagacgccatccctgaagagagtgagatg 1296
Db 1190 CACCTGGGGTCCAAATGACGAGCGATTCCTGAGAGCCATCCCTGAGGAGAGTGGCGATG 1249
Qy 1297 agagcaagaagacaccttgcaagcgaacttcgaacttcgaacttcgaacttcgaacttcga 1356
Db 1250 AGAGCAAGACCAACCTGACAGCGATCTGATGCTGCTCTGCAAAACGAAATGCTT 1309
Qy 1357 gtgaggaagaagttctcgaacttcgaagagagagagagagagagagagagagagagagagag 1416
Db 1310 GAGAGCAAGAGTTCTCCGATTCGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1369
Qy 1417 acttcaaaaagcgaag 1476
Db 1370 ACTTCAAAAAGCCAG 1429
Qy 1477 agaaagaagatccagcag 1536
Db 1430 AGAAGAAAGTCAACCGAG 1489
Qy 1537 agagagagagcgaagtgagcgaacttcgaacttcgaacttcgaacttcgaacttcgaacttc 1596
Db 1490 AGAGAGAGGTCAAGTTGGGCTGATGAGTACCTCTCCAGCTCTGCTGCTGCTGCTGCTGCT 1549
Qy 1597 cagcttcttccc 1610
Db 1550 CACGTTTCTTCCC 1563

RESULT 5

LOCUS ARO12007 2111 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 2 from patent US 5763182.
ACCESSION ARO12007
VERSION ARO12007.1 GI:3969997
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2111)
TITLE Nakamura, Y. and Furukawa, Y.
RPPD. protein and DNA encoding the same
JOURNAL Patent: US 5763182-A 2 09-JUN-1998;
FEATURES
Location/Qualifiers
1..2111
Source /organism="unknown"

BASE COUNT 572 a 485 c 544 g 510 t
ORIGIN

Query Match 93.0% Score 1498; DB 6; Length 2111;
Best Local Similarity 99.3% Pred. No. 0;
Matches 1504; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

maybe probe

audience

Qy 97 gggagagcgaagaatgagcgaagaagcgaagcgaagcgaagcgaagcgaagcgaagcgaagcga 156
Db 50 GGGAGGCGGCAAGATGGCGAGCGACGACCGCGGAGAAAGTCTGTTACTACTACG 109
Qy 157 acggagatcttgaaattacttattgacagcgaacccaatgaagctcaccgaatcc 216
Db 110 ACGGGATGTTGGAAATTAATTTATGACAAAGGCCACCAATGAGCCTCAACCGATCC 169
Qy 217 gcatgactaataatctgctcgaactatgctcctacgaaataatgaaatcgaatcgc 276
Db 170 GCATGACTCAATATTTGCTGCTCAACTATGCTCTACGGAATAATGAAATCTATGCGC 229
Qy 277 ctcaacaagaagcgaatctgagagagatgacccaagtaaccaagcgaatgaactaataattct 336

Db 230 CTCACAAACCAATGCTGAGAGATGACCAAGTACACAGCATGACTCAATTAATTCT 289
Qy 337 tgggtcccaatccgtcagaatacaatgagagatgacgaagaagatgacgaagaatccaag 396
Db 290 TGGGCTCCATCCGTCAGATTAACATGTGGAATGACGAAAGAGATGCAATTCACAG 349
Qy 397 ttggtgagagcgtccagatccgatggcctgtttgagtcctgacagtgctactggtg 456
Db 350 TTGGTGAAGACGTCCAGATTCGATGAGCGCTGTTGAGTTCGTGCTGCTGCTGCTGCTG 409
Qy 457 gtctgtgcaagtgctgtgaacttaataagcagaagacggaatcgcgctgaaattggg 516
Db 410 GTTCTGTGCAAGTGGTGAAGCTTAATAACAGACGACGACATCCGTGTGAATTGGG 469
Qy 517 ctgggggacctgacccaatgaagaagaagtcgaggaacttcgtcttcgtgaactgaatga 576
Db 470 CTGGGGGCTGCAACCTGCAAAAGAGTCCGAGGCACTTGCTTGTGTAAGTCAATGATA 529
Qy 577 tgccttgccatccctggaactgctaagatcaccaagaagggtgctgtacatgacatg 636
Db 530 TCGTGTGGCCATTCCTGGAACCTGTAAGTATCACAGAGGGTGTGATTCATTCATTTG 589
Qy 637 atattcaaccaatgagtgagcgttggaagagccttcacacgaagccgggtcatgactg 696
Db 590 AATATTCACCATGATGACGCGCTGGAAGAGCCCTTGTACACGACGACGCGGTCAATGACTG 649
Qy 697 tgccttcaatagatgag 756
Db 650 TGTCTCTTATATATATGATGAGAGATCTTCCAGAACTGGGAGCTCACGGGATATCGGGG 709
Qy 757 ctggcaaaagacaagatattatgctgttaactacccgtcccgaaagcggatgtgaatga 816
Db 710 CTGGCAAAAGCAAGATATATGCTGTAACTACCGCTCCGACAGCGGATATGATGAGAGT 769
Qy 817 cctatgagagcatttcaagccggtcaatgltccaagtaatgagatgttcaagcctagt 876
Db 770 CCTATGAGGCCAATTTTCAAGCCGGTCAATGTCCAAAGTAAAGATGTTCCAGCCTAGTG 829
Qy 877 cgggtgcttcaagtggtgag 936
Db 830 CGGTGCTTTACAGTGTGCTCAGACTCCCTATCTGGGAGTGGTAAAGTCTTCAATC 889
Qy 937 taactataaagaagacagcgaagtgtgtgaatttgtcaagaagcttaactcgtcactatgc 996
Db 890 TAATATATCAAAAGACACGCCAAGTGTGTGAATTTGTCAAGAGCTTAACTGCTTATGC 949
Qy 997 tgaatctggagagcgtgtgttacacattcgtaaacgttgccgggtgctggaatataatga 1056
Db 950 TGAATCTGGAGGCGGTGATACACCATTCGTAAACGTTGCCGCTGACAGGACATATGAGA 1009
Qy 1057 cagcgtgccccgtgataagagagatcccaatgaagcttcacataactgaactgaat 1116
Db 1010 CAGCTGTGGCCTTGATAGAGAGATCCCAATGAGCTTCATACATGACTCTTTGAT 1069
Qy 1117 actttgaccagattccaagctcacatcagtccttccaataatgaactaaccagaacaga 1176
Db 1070 ACTTTGGACCAATTTCAAGCTCCACATCAGCTTCCAAATATGACTTAACAGAACAGCA 1129
Qy 1177 atgagtaacttgagaagaatcaaacagcgaactgtttgagaaccttaagaacttcgcgaacg 1236
Db 1130 ATGAGTACTGGAGAGATCAAAACAGCACTGTTTGAAGAACTTGAATGCTGCGCAGC 1189
Qy 1237 caacttgggttccaacacgcagcagatctccttgaagacgccatccctgaagagagtgagatg 1296
Db 1190 CACCTGGGGTCCAAATGACAGCGATTCCTGAGAGCCATCCCTGAGGAGAGTGGCGATG 1249
Qy 1297 agagcaagaagacaccttgcaagcgaacttcgaacttcgaacttcgaacttcgaacttcga 1356
Db 1250 AGAGCAAGACCAACCTGACAGCGATCTGATGCTGCTCTGCAAAACGAAATGCTT 1309
Qy 1357 gtgaggaagaagttctcgaacttcgaagagagagagagagagagagagagagagagagagag 1416

Db	1310	GTGAGGAAGAGTTCCTCCGATTTCTGTAAAGAGAGGAGAGGGGGGCCCAAGAACTTTTCCA	1369
Qy	1417	acttcaaaaaaagccaagsgagatc caaacacagagat tgaataaagaagaagaccagaagaga	1476
Db	1370	ACTTTAAAAAACCACAGAGAGTCAAAACAGAGATGAAAAAGAAACCCAGAGGAGA	1429
Qy	1477	agaaaggaatcacccaagagaggaataacccaagagaggaagc cagaagccaagaaggtgtca	1536
Db	1430	AGAAAGAAGTCAACCACAAGAGGAGAAACCAAGAGAGAGACCCAGAACCAAAAGGGGTCA	1489
Qy	1537	aggagagagccaagltggtgc tgaat tgaacctctcagctctgtgctctgtgtgagtcct	1596
Db	1490	AGGAGAGAGTCAAGTGGCGCTGAATGAGACCTCTCCAGCTGTGGCTTCTGTGATGATCCCT	1549
Qy	1597	cacgtttcttccc 1610	
Db	1550	CACGTTTCTTCCCC 1563	

RESULT	6				
Accession	162388		2111 bp	DNA	
Version	162388				
Accession	162388				
Keywords	162388.1		GI:2480336		
Source	Unknown.				
Organism	Unknown.				
Reference	Unclassified.				
Authors	1 (bases 1 to 2111)				
Title	Nakamura, Y. and Furukawa, Y.				
Journal	RPDL protein and DNA encoding the same				
Features	Patent: US 5659016-A 2 19-AUG-1997;				
Source	Location/Qualifiers				
	1. 2111				
	/organism="unknown"				
Base Count	572 a		485 c	544 g	510 t
Origin					

no reference

we have prepared

primers

Query Match	93.0%	Score 1498	DB 6	Length 211
Best Local Similarity	99.3%	Pred. No. 0		
Matches 1504	Conservative	0	Mismatches 10	Indels 0
				Gaps 0

OY	97	ggagagcagcaagaatggcgacagacagagagaccccgagagaatctgttaactacg	156
Db	50	ggagagccagcagaatggcgcgacagcgacgacccggagaaaattctgttactacg	109
OY	157	acggagatgttggaaattactatlaaggacaagccaccccaatgaagctcaccgaatcc	216
Db	110	acggggagattggaaattactattttatggacaagcgaccaccacccaatgaagctccaccgaatcc	169
OY	217	gcataccataatttgtctccaacttgytctctacgcgaaaaatggaaatctctgc	276
Db	170	gcatactatattattttgtcgtccaacttattgtctcttaccgaaaaatggaaattcttgcgc	229
OY	277	ctcaacaagccaatgctgaagatgagacaagtaaccacagcgatctactataatct	336
Db	230	ctcacaaagccaattgctgagagatgaccacaattaccacagcattactatcatttaattct	289
OY	337	tgcgctccatccgctcagataacaatgctggagttacagacaagacagatgcagagatccaag	396
Db	290	tgcgctccatccgctcagataacaatgctggagttacagacaagacagatgcagaaattccaag	349
OY	397	ttggfsgaagacgtccagatattcagttgscgttttgagttctgtcaagttgtctactggtg	456
Db	350	ttggfsgaagacgtccagatattcagttgscgttttgagttctgtcaagttgtctactggtg	409
OY	457	gtctctgtggcaagtctctgaaactaaataagcaagacagagacatcgccgtgaaattgg	516
Db	410	gtttctgtggcaagtctctgaaacttaattaaacacacagacgacactccgttgaaattggg	469
OY	517	ctgggggacctgcacatgcagaaagaatccgagatctgtgcttctgttaagctaaatgata	576

Db	470	CTGGGGGCGCTGCACCATGCAAGAAATCCGAGGCATCTGGCTTCTGTACTCATATGATA	529
Qy	577	tcgctcttgccatcccttggaactgctaaagtatccacagaggtgctgtacatgtacatg	636
Db	530	TGCGCTTGGGCATCTCGGAACCTGCTAAAGTATACACAGAGGTGCTGTACATTGACATWG	589
Qy	637	atatacaccatgttgacgsggttgaaagaagcctcttaaccagagaccgggtacatgt	696
Db	590	ATATTACACCATGGGAGGGGTGGAAAGAGCCCTTCTACACACGAGACCGGGGTCAATGATG	649
Qy	697	tgctcccttcaaaagtatgagagtagtctccagagacatgaggaccctaaaggtatccaggg	756
Db	650	TGTCTTTCATTAAAGTATGAGAGTACTTCCCGAAGAACTGGGAGCTTACGGGATATCGGGG	709
Qy	757	ctggcaagaacaaagtattatgtctgttaactaccgcgtccgagacggatgtgtagaagt	816
Db	710	CTGGCAAAAGGCAAGTATTATAGCTGTTAACCTACCCGCTCCGACACGGATTTATACAGAGT	769
Qy	817	ccctatgggcacatttcoaagccggtcatgtgccaaagtaaigtatccagcctagt	876
Db	770	CCATTGAGGCGCATTTTTCACCCGCTATGTCCAAAGTATGAGATGTTTCAAGCCTTATGC	829
Qy	877	cgggtgtcttaacagtgctgtcctagactccctatctctggagatcgglttaagttgtctaac	936
Db	830	CGGGGTCTTACAGTGTGGCTCAGACTCCCTTATCGGGGATCGGTATGATTGTCTTCAATC	889
Qy	937	tatctacaaaggacaagcccaagtggtgtgaaattgttcaaaagccttaactgtcctatgc	996
Db	890	TAACTATCAAAAGGACACGCCAAAGTGTGTGAATTTGTCAAGAGCTTTTAACTGCTTATGC	949
Qy	997	tgaatcttggaaggcgggtggtgtatcaccaatcttgtaacglttgcccggtgtctgacatatga	1056
Db	950	TGATGCTGGGAGGGGGGTGTATCACCACTTCGTAAAGTTGCCGGGTGCGAGCAATATGAGA	1009
Qy	1057	caggtgtggcccttggtatcgggaatctccaaatgagcttccataacatbaactaacttgaat	1116
Db	1010	CAGTGTGGCCCTGTGGATACGGAGATACCTTAATGAGCTTCCATACATATACTACTTTGAT	1069
Qy	1117	actcttgacccaagatttcaaagctccacatagcttccttccaatatgtactaacacagaaca	1176
Db	1070	ACTTTGGACCCAGATTTCAAAGCTCCACATCAGTCCATCTCCCAATATGACTAAACCAAGAA	1129
Qy	1177	atgaatctctggagagaagatacacaagcgagactgtttgagaaccttaagatgtctgcgcag	1236
Db	1130	ATGAGTACCTCGGAGAAAGATCAAAACACGCACTGTTTGAGAACTTGTAAGTGTGCCGACG	1189
Qy	1237	caacttggtgtccaaacgcaagcgatctccgtgagaagccatccctgaagagaagtgtagatg	1296
Db	1190	CACCTGGGGTCCAAATGCAAGGCATTTCTGTAGAGACGCCATCCCTGTAGAGAGAGTGGCATG	1249
Qy	1297	agagcaagaagacccctgacaagcgacatctgatactgtctctctgtacaacgaattgact	1356
Db	1250	AGGACGAGAGACGACCCCTCGAACACCGCATCTGCATCTGCTCTGTACAAAGAAATTGGCT	1309
Qy	1357	gtggagagaaggtcttccgattcttgaaagagagagagagagggggcgcaagaacatcttcca	1416
Db	1310	GTGAGGAGAGATTCTCCGATTCTGAAAGAGAGAGAGAGGGGGCCGCAAGAACTTTTCCA	1369
Qy	1417	acttcaaaaaagccaagaggtgtlcaaaaacagagagatgtaaaagaagaaagaccacagagaga	1476
Db	1370	ACTTCAAAAAAGCCAAAGAGTAGTCAAAAACAGAGAGTAAAAAAGAAACCCACAGAGAGA	1429
Qy	1477	agaaagaataccaccgaaagagagagaaaaacaaagggagagaaagcccaagcaaaaggtgtca	1536
Db	1430	AGAAAGAAGTATCCAGGAAGAGGAAGAAAAACCAAGAGAGAGAGGACCAAAACCCAAAGGGGTCA	1489
Qy	1537	agagagagcccaagttggaccttgatagactctccaaagctctgctctctctgtgagtcct	1596
Db	1490	AGGAGAGAGGTCAAGTTGGCTCTGAATGAGACTCTCCAGAGCTGTGGCTTCTGCTGAGTCCCT	1549
Qy	1597	caagttctcttccc 1610	

Dn	1550	CACGTTCTTCC	1563
RESULT	7		
ID	E11455	standard; RNA; HMW; 2111 BP.	
XX	E11455;		
AC	E11455;		
SV	E11455.1		
DT	07-OCT-1997 (Rel. 52, Created)		
DT	02-SEP-2000 (Rel. 65, Last updated, Version 2)		
DE	cDNA encoding human transcriptional control protein, RPDL.		
XX	JP 1996140687-A/1.		
XX			
OS	Homo sapiens (human)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;		
OC	Eutheria; Primates; Catarrhini; Homiidae; Homo.		
XX	[1]		
RP	1-2111		
RA	Nakamura Y., Furukawa Y.;		
RT	"RPDL PROTEIN AND DNA ENCODING FOR THE SAME";		
RL	Patent number JP1996140687-A/1, 04-JUN-1996.		
RL	JAPAN FOUND CANCER RES, EISAI CO LTD,.		
XX			
OS	Homo sapiens (human)		
PN	JP 1996140687-A/1		
CC	PD 04-JUN-1996		
CC	PF 20-JUL-1995 JP 1995183763		
CC	PR 22-SEP-1994 JP 94P 1227876		
CC	PI NAKAMURA YUSUKE, FURUKAWA YOICHI		
CC	C12N15/09,C07K14/39,C07K16/18,C12P21/02,C12P21/08,C12Q1/68,		
CC	PC C12P21/02,		
CC	C12R1:19);		
CC	PC strandedness: Double;		
CC	CC topology: Linear;		
CC	CC hypothetical: No;		
CC	CC anti-sense: No;		
CC	FH Key Location/Qualifiers		
CC	FH source		
CC	FT 1..2111		
CC	FT /organism="Homo sapiens"		
CC	FT /tissue-type="Lung"		
CC	FT /clone="L1-3977"		
CC	FT 1..63		
CC	FT CDS		
CC	FT 64..1512		
CC	FT /product="human transcriptional control		
CC	FT protein,RPDL"		
CC	FT 1513..2111		
CC	XX key Location/Qualifiers		
EH	source		
EH	FT 1..2111		
EH	FT /db_xref="taxon:9606"		
EH	FT /organism="Homo sapiens"		
SQ	Sequence 2111 BP; 572 A; 485 C; 544 G; 510 T; 0 other;		
Query Match	93.0%	Score 1498; DB 23; Length 2111;	
Best Local Similarity	99.3%	Pred. No. 0;	
Matches 1504; Conservative	0; Mismatches	10; Indels	0; Gaps
OY	97	ggagaggcagaagaatgagcgacagcgaagccgcgaggaagtctgttactactagc	156
DB	50	GGGAGGCGAGCAAGAATGGCCGACGACGAGGACC CGAGGAAGTCGTACTACTAGC	109
OY	157	acggggatgttggaattactatatgagcaaggccaaccatgaagcctcacccaatcc	216

D	b	110	ACGGGAGTGTGGAAATTA	CTATTATTA	GCACAAAGGCCACCCATTGA	AGGCTCACGSAATCC	169
Q	y	217	gcatactcaataat	ttgtcgtctcaactat	tgctctcaccgaaata	tgaaatctatgcgc	276
D	b	170	GCATGACTCATAAT	TTTGTCTCTCACTAT	TGTCTACACGAAATGAAATCT	TTCTGCCC	229
Q	y	277	ctcgaagccaat	tgctcgtgaggaatga	ccaagttaccacagcga	tgactcacattaatctc	336
D	b	230	CTCCAAAGGCCAAT	GTCTGAGAGATAT	CCAGATACACACGCGATACATCA	TTAAATTTCT	289
Q	y	337	tgagctccatcgc	ctccgaataacat	ctgagatacgcgaacagata	tgcaagatctcaacg	396
D	b	290	TGGCCTCATCT	CGTCCAGATACAT	TGTCCGAGTACACGACAGATG	CAGATTCACACG	349
Q	y	397	ttggtgagagac	ctgtccagialtcg	atgtagccctgtttgagttc	gtcagttgtctactggtg	456
D	b	350	TTGTGTAGGACTG	TCAGTGTTCGATGTGGCCTG	TTTGTGATTTCTGTCA	TTGTCTACTGTG	409
Q	y	457	gtctctgtgc	aaagtgtcgtgaaact	taataagcagcagacggaact	tcgcgtgtaattggg	516
D	b	410	GTTCGTGGCAAG	CGTGTCTAAACTTAA	TAAAGAGCAGACGACATCCCT	GTGAATTTGGG	469
Q	y	517	cttgaggagcc	ctgcacacatgca	aaagaatccgagagcat	ctgtcgtctcgtatc	576
D	b	470	CTGGGGGCTGTG	CACTACATGCAAAAGAT	TCGAGGCACTGTGCTTCTG	TTATGCAATGTGA	529
Q	y	577	tcgactctgag	catccctctggaact	ctgtaaaatlatccacagaggt	tgctgtacatltgacat	636
D	b	530	TCTGTTTGGCAT	CTCTGCGATCTG	CTAAGATACACAGAGGGTG	CTGTATGTGACATTTG	589
Q	y	637	ataatccaat	tgatgacgcgcgt	tggaagaggtcctctacacagc	gaacggtatctgactg	696
D	b	590	ATATTTACCACT	GTGGTAGCGGTGTGAAG	AGCGCTTCTACACACGAGACCG	GGGTCAATGACTG	649
Q	y	697	tgctccctca	taagiatatgagag	tgacttccacagaaact	tgaggactacggataccggg	756
D	b	650	TGTCTTTCTTA	AGTATGAGAGAGT	ACTTCTCCAGAAACT	TGGGACCTACGGGATATCGGG	709
Q	y	757	cttgcaaaaga	caaglatlatat	tgctgttctaactaccgcgt	ccgagacggatltgat	816
D	b	710	CTGSCAAAGCA	AGTATTATTTAGCTG	TTAACTACCCGCTCCGAGACGG	ATTTGATGACGAGT	769
Q	y	817	ccctatgag	gccatlttccaag	ccggtcatatgtccaag	taatgtagatgtgtcca	876
D	b	770	CCATTGAGGGC	ATTTCACCCGGTCA	TGATGTCACAAATTA	TGAGATGTTCAACGCTTAATG	829
Q	y	877	cggtggtcct	taagatgtgtgct	tcagactccatctc	tgagatcgtgttaagt	936
D	b	830	CGGTGCTTAC	AGTGTGTGCTCAG	ACTCTCTATCT	TGGGATGTGGTTAGTTC	889
Q	y	937	latcatcaaa	agagacacgccc	aagtctgtgtgaa	tttgtcaagaagctt	996
D	b	890	TAACTATCA	AAAGGACACAGCC	ACAGTGTGTGAATTTGT	CAAGAGCTTTAACTG	949
Q	y	997	tgtatgct	cggaagcgtgtgt	gtltaacatctcg	taagcttgcccggtgtc	1056
D	b	950	TGAATGCTG	GAGGAGCGGTGT	TACACCATTCGTGA	ACCTTCCGGGTGCAGGAC	1009
Q	y	1057	caagtgtgag	ccctcgataagagagat	ccctcaatbagctt	ccataaataataact	1116
D	b	1010	CAGTGTGTG	CCCTCTGATACG	GAGATCTCCCTTA	TATGAGCTTCCATACATG	1059
Q	y	1117	actttgga	ccaagatttcaagct	ctccacatagctctt	ccataatgatacca	1176
D	b	1070	ACTTTGGAC	CAAGATTTCCAAAGCT	CTCACATCAGTCTTCC	CAATATGTACCTAAC	1129
Q	y	1177	atgagta	ccctgsgaagat	atcaaacagcgcgc	gtgttgagaaacct	1236
D	b	1130	ATGAGTACT	CTGGAGAAAGATCA	AAACACGCACTGTT	TGAGAACTTTGATATG	1189
Q	y	1237	caactgtg	gtctcaaaogcagag	gcgattccctgtgagagc	gcatacccttgtaggaag	1286
D	b	1190	CACCTGGGGT	CAAAATTCACAGCC	CAATTCCTGAGAGCC	CACTTCCCTTAGAGAGT	1249

QY 1297 aggaagacacccctgacacgcatctgctctctctgacacgaatctgct 1356
1250 AGGACGAGACACCTGACAAAGCATCTGATCTCTCTGACAAAGCATCTGCT 1309
QY 1357 gtgaggaagatctctgacatctgagagagagagagagagagagagag 1416
1310 GTGAGGAAGATCTCTGATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1369
QY 1417 actcaaaaagacaaagagatcaacaagagatgataaagagagagagagag 1476
1370 ACTTCAAAAAGACCAAGAGATCAACAAGAGATGATAAAGAGAGAGAGAGAG 1429
QY 1477 agaaaggaacacagagagagagagagagagagagagagagagagag 1536
1430 AGAAAGAGATCACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1489
QY 1537 aggaag 1596
1490 AGGAGAGAGATCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1549
1597 caggttctctcc 1610
Db 1550 CACGTTCTCTCC 1563

RESULT 8
LOCUS MMHIDE 1977 bp mRNA linear ROD 09-SEP-1997
DEFINITION M. musculus mRNA for histone deacetylase.
ACCESSION X98207
VERSION X98207.1 GI:1771285
KEYWORDS histone deacetylase.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1977)
AUTHORS Bartl,S., Taplick,U., Lager,G., Khler,H., Kuchler,K. and Seiser,C.
TITLE Identification of mouse histone deacetylase 1 as a growth
factor-inducible gene
Mol. Cell. Biol. 17 (9), 5033-5043 (1997)
JOURNAL 97415582
MEDLINE 2 (bases 1 to 1977)
REFERENCE Seiser,C.
AUTHORS Direct Submission
TITLE Submitted (30-MAY-1996) C. Seiser, University of Vienna, Institute
JOURNAL of Molecular Biology, Vienna Biocenter, Dr. Bohr-Gasse 9, A-1030
Vienna, AUSTRIA

FEATURES

source

1. 1977
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="Swiss 3T3"
18..1466
/codon_start=1
/product="histone deacetylase"
/protein_id="CA66870.1"
/db_xref="GI:1771286"
/db_xref="MGD:MGI:108086"
/db_xref="SWISS-PROT:O09106"
/translation="MAOTGKRVVCTYYDGVNYYGGHPKPHIRMTNILLN
YGLRKMEIYRPHKANAEMTKYHSDYIFLRSIRPDNMSYKQMRFWGDCPV
FDGLFEFCOLSTGSGVASAVLKNQODIANNMGJHAKKSSAFSCYNDIVLAI
LELTKHORYLYIDIDHGDVGEAFYTDVRVMTVSFHKXGEYFGDGLRDIGAK
KXYAVNYPLRDGIDDESYSYAIKPKVSKYEMOPASVYLCSDSISGRDLCFNL
TKGHAKCYEPYKFNPLMLGSGGTTIRNVAWCWYETRAVALDTENLPLPYNDP
ETRPDLKHSNMTNONTNELYELIKOLFEFLNRLPAAPGVOMALPEDAIPPE
SGDEDEDPKRISICSSDKRIACEEESDSDEGGRKSNFKKAKRVKTEDEKE
KDPKEKEVEEETKEKEPEAKVKEVKLA"

CDS

BASE COUNT 543 a 475 c 488 g 471 t
ORIGIN

Query Match 77.0%; Score 1240.2; DB 10; Length 1977;
Best Local Similarity 89.2%; Pred. No. 0;
Matches 1349; Conservative 0; Mismatches 163; Indels 1; Gaps 1;

QY 94 gcgaggagcgagcaagatgagcgagagagagagagagagagagagag 153
1 GCCGGGCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 154 acgagggagatgtgaaatctatctatgacaaagggcccaatgaaagcga 213
61 ACGAGGGGATTTGGAAACTCTATTATGACAAAGGGACCCCATGAAAGCTCA 120
QY 214 tccgagactcaataattgtctgaactatgctctcaccgaaatgaaatc 273
121 TCCGATGACTCACAATTGCTGCTCAACTATGCTCTCAACGAAATGAGATCT 180
QY 274 gccctcaaaaagcaatgctgagagagagagagagagagagagagagag 333
181 GCCCTCACAAAGCCAAATGCTGAGAGATGACCAAGTACCAAGTATGACTA 240
QY 334 tcttgagctcaatccgctcagaatacatgctcgagatagacagagagatga 393
241 TCTTGAGCTCTATTGCGCCAGATTAACATGCTGATTAACAGCAAGCATGCA 300
QY 394 acgttgagagagactgctcagatcagatgagcgtgttgagttcgtcagtt 453
301 ATGTGAGTGAAGAGCTGCTCGGATTTGAGGCTTTGAGTGTGCTGCTCAGG 360
QY 454 ggtgtctgtgagagagagagagagagagagagagagagagagagag 513
361 GAGGCTCTGTCGAAGTGTGTAAGCTTAATTAAGAGAGAGAGAGAGAGAG 420
QY 514 gggctgagggcctgacacacacacacacacacacacacacacacacac 573
421 GGGCTGAGGGGCGCTGCACCATGCAAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 574 atatgctctgagacacacacacacacacacacacacacacacacacac 633
481 ACATGCTCTGAGCCATCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 634 ttgatattacacatgtgagagagagagagagagagagagagagagag 693
541 TTGATATTACCATGAGCGAGTGGCGTGAGAGAGAGAGAGAGAGAGAGAG 600
QY 694 ctgtctctctcataagtaagtagagagagagagagagagagagagagag 753
601 CGTGCTCTTTCATTAATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 754 gggctgagcaagagacagatattgctgttaactaccgcgtccgagagag 813
661 GGGCTGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 814 agctctatgagagcatttcaagccggtcatgttccaaagtaatgagatgt 873
721 AATCTATGAAAGCCATCTTAAGCCAGTATGTCCAAAGTAATGAGAGATTC 780
QY 874 gtgcgagtgcttaagtagtgctcagagctccatctcggagatgagtagt 933
781 GTCAGAGTGTCTTACAGATGTGCTCAGATTCCTGTCTGGGAGCGGTTAG 840
QY 934 atctatctataaagagacagcgaagtgtgaaattgttcaagaagcttaac 993
841 AATCTAGCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 994 tgcgtatgtcgtgagagagagagagagagagagagagagagagagag 1053
901 TGCATATGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 1054 agaaagcgtgtgagcctgtgagagagagagagagagagagagagagag 1113
961 AAACAGCGGTGCGCTTGAGACAGAGATCCCTAATGAGCTGCCCTTACAT 1020

QY 1114 aatacttgaccagatttcaacgtccacatcagcttccaatatgactaacagaaca 1173
 Db 1021 AATATTGGACCGGATTTCAAGCTTCAATCAGCCCTTCAATCAATGACCAACCGAACA 1080
 QY 1174 cgaatgtaaccttgagaagaatcaacaacgagctgtttgagaaccttagatgtgcgc 1233
 Db 1081 CTAAAGAGTACTCTGAGAGATCAAGACGCTCTTTCAGAACTTGAGATGATGCCCC 1140
 QY 1234 acgacacgtgggttccaaacgcagaggtctctcctgaagacgacccctgaagagatggcg 1293
 Db 1141 ATGCCCCCTGGGCTCCAGATGCAAGGCCATCCTTGAGAGCGCATCCCGAGAGAGTGGGG 1200
 QY 1294 atgaagaaagaaagcagaccttgcaagcgatcgcgtctcctctgacaacgaatly 1353
 Db 1201 ATGAAGATGAGGAGGAGCCTGACAAACGATCTCATCTGCTCTGTATTAACGATTTG 1260
 QY 1354 cctgtgagaaagatgtctccgacttctgaagagagagagaggggcgcaagaactct 1413
 Db 1261 CCTGTGAGGAAAGATTCCTGCGACTCAGATGAGAGGAGGAAAGTGTGCGCAAGACTCTT 1320
 QY 1414 ccaacttcaaaaagcgaagagatcaaaaagagagatgaagaagaagagagagag 1473
 Db 1321 CTAACTTAAAAAAGCCCAAGAGTTAAACAGAGGATGAGAAAGAGATCCTGAGG 1380
 QY 1474 agaaagaaagaaatccagaaagagagagaaacccaagagagagagagagagag 1533
 Db 1381 AGAAAAAGAGTACAGAGAGAGAGAGAAACCAAGAGAGAGAGAGAGAGAGAGAGAG 1440
 QY 1534 tcaag 1593
 Db 1441 TCAGAGAGAGAGTCAAGTGTGCGT-AGCAAGTCTGAGAGGCCCATCTTCCCAAGTT 1499
 QY 1594 cctcagcttctt 1606
 Db 1500 CCTCACTCTCCT 1512

RESULT 9
 MMU80780 1999 bp mRNA linear ROD 02-SEP-1997
 LOCUS Mus musculus putative histone deacetylase (HDI) mRNA, partial cds.
 DEFINITION U80780
 ACCESSION U80780.1 GI:2347179
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Need

EREFERENCE 1 (bases 1 to 1999)
 AUTHORS Johnson, C.A.
 TITLE Putative histone deacetylase HD1 from M. musculus
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1999)
 AUTHORS Johnson, C.A., White, D., O'Neill, L.P. and Turner, B.M.
 TITLE Direct Submission
 JOURNAL Submitted (02-DEC-1996) Department of Anatomy, University of
 Birmingham, Birmingham B15 2TT, U.K.
 COMMENT On Sep 2, 1997 this sequence version replaced gi:1737470.
 FEATURES
 location/Qualifiers
 1..1999
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 1..1999
 /gene="HDI"
 /gene="HDI1"
 15..1463
 /note="similar to human histone deacetylase HD1 and yeast
 RPD3"
 /codon_start=1
 /product="putative histone deacetylase"
 /protein_id="AAB68398.1"
 /db_xref="gi:2347180"
 /translation="MAQTGTRRVYVDGNGVYGGCHPMXPHRIRMTNHLIN

CDS

/organism="Mus musculus"
 /db_xref="taxon:10090"
 1..1999
 /gene="HDI"
 /gene="HDI1"
 15..1463
 /note="similar to human histone deacetylase HD1 and yeast
 RPD3"
 /codon_start=1
 /product="putative histone deacetylase"
 /protein_id="AAB68398.1"
 /db_xref="gi:2347180"
 /translation="MAQTGTRRVYVDGNGVYGGCHPMXPHRIRMTNHLIN

BASE COUNT 563 a 467 c 480 g 467 t 2 others
 ORIGIN

Query Match 76.7%; Score 1236.4; DB 10; Length 1999;
 Best Local Similarity 89.0%; Pred. No. 0;
 Matches 1344; Conservative 2; Mismatches 163; Indels 1; Gaps 1;

YGLYRKEIYRPHKANAEEMTKYHSDYIKFLRSIRPDNNSEYKQMRNVEDCPY
 FDGLPEFCOLSTGGSVASVAKLNOQDIDIANWAGJLHAKSASGECYNDVLAI
 LELKYNQRLVYIDIDHHGQVREARITDNRVTFSEHKGSEFPEPTGDLRGAK
 GKTYAVNPPLRDGIDDESIEYERKPVSKVHEMOPSAVVIQGSDDLSDGRLCFNI
 TIKHACVEFKSFNPLMLGGGCTTIRVACWYETRAVALDTEIPLELPNDYF
 EYGPDEKRLHISPNMTNONTNEXLEIKORLPENLMLPAPVQMOALPEDAIPEE
 SGDEDEPDKRISICSDKRIACEEFSDEEGEGRKNSSNFKRAKRVKTEDEKE
 KDPEKKEVTEERKEEKEPRAKVEVKLA"
 QY 97 gggagagcagcaagaatgagcgcagacgcagggaccccgagaagaatctgttactactcg 156
 Db 1 GGGCGCGCAGCAAGATGCGCGCAGACTCAGGCGCACCAGAGGAAAGTCTTACTACTACG 60
 QY 157 acggagatgttgaataattactatattgacaagagcccaaatgaagctcaccgaatcc 216
 Db 61 ACGGGGATGTTGGAAACTACTATTATGACAGAGGCGACCCCATGARKGCTCACCGAATTC 120
 QY 217 gcatctacataattgtctgtcacaactatggtctctacagaaaatggaaatctatcgcc 276
 Db 121 GCATGACTCACAAATTTGCTGCTCAACTATGGTCTCTACCGAAATATGAGATCTCACGTC 180
 QY 277 ctcaaaaagccaatgtctgagagagatgacccaagttaccaagagatgactacataatct 336
 Db 181 CTCACAAAGCCAAATCCTGAGGAGATGACCAAGTACCAAGTACATACATTAATTC 240
 QY 337 tgcgtccatccgtccagataacatgctcgagatgacgaagcagaatgagagatcaag 396
 Db 241 TGCGTTCTATTGCGCCCAATAATCAATGCTGATACAGACAGATGAGATTCATG 300
 QY 397 ttgtgtgagactgtccaglatcgatgagcctgttgaattctgtcactgtctactggtg 456
 Db 301 TTGCTGAGAGAGACTGTCGGTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 360
 QY 457 gtctgtgtgagactgtctgtgaacttaataagcagaagagactgtgcgtgaattgg 516
 Db 361 GCTGTGTGCAAGTGTCTGAGAGCTTAATTAACACACACACAGGACATCTGTGAATCTGG 420
 QY 517 ctgggggctgcaccatgcgaagaagtcgagagcactgtgcttctgttctacatgata 576
 Db 421 CTGGGGGCTGACCATGCAAGAGAGTGTGAAGCATCCGGCTTGCTGTACGTCATGACA 480
 QY 577 tgcgtctgtgcaatccctgtgaactgtctaaaglatcaccagaaggtgtctgtacatgacat 636
 Db 481 TCGTCTTGCCCATCCTGGAACGTCTAAAGTACACACAGAGGAGTGTCTCTATTTGACATTTG 540
 QY 637 atattaccatgtgtgagcgtgtgagaagagccttctacacagagacgggtatgactg 696
 Db 541 AATATCACCATGGCGATGCGGTGAAGAGCGCTTCTATACACACCGGGCATGACTG 600
 QY 697 tgccttccataagatgagagagacttcccaagagactggggagacctgaagataccgggg 756
 Db 601 TGCTCTTTCATTAATAGCGAGAGTACTCTCCAGAGACGCGGAGACTACGGAGCATTTGGGG 660
 QY 757 ctggcaagaagaatataatgtctgttaactaccgcgtccgcgaagcggagttgatgagagt 816
 Db 661 CTGGCAAGGCAAGTACTATGCTGTGAGACTACCCACTCGAGAGCGCATTTGACGACGAT 720
 QY 817 cctatgagccatlttcaagcgcgtgtcacaaggtataatgagatgttccagctagt 876
 Db 721 CCTATGAAGCCATCTTTAAGCCAGATGATCCAAAGTATGAGATGTTCCAGGCTATG 780
 QY 877 cgggtgttcaaggtgtgtgcagaactccatctcggggagacgggttaagttgttcaatc 936
 Db 781 CAGTGTCTTACAGTGTGTGCTCAGATTCCTGTCTGGGGACCGGTAGGTTGCTTCAATC 840
 QY 937 tatcatcaagagacagcgaagtgtgtgaattgttcaagaagcttcaactgtccatgc 996

```

Db      841 TGACCATTAAGAGACACGCAAGTGTGTGAGATTGTGAAGATTCAACTTGCCCATGC 900
Qy      997 tgaatctggagagcggtgtgttaccatctgtaacgttgcgcgtgctgagacataatga 1056
Db      901 TGATCGCTGGAGAGAGAGTGGCTACACATCCGGAATGTTGCTGCGCTGAGATTGCGAAA 960
Qy      1057 cagctgtgcccctgtaacagagatccctaataatgagcttcacatacagactcttgaat 1116
Db      961 CAGCGGTGGCCCTGGACACAGATCCCTAATGATGCTGCCCTACAAAGCACTACTTGAAT 1020
Qy      1117 acttgagacagatttaagctcccatcagctcctcaatgataatgatacagaaacaga 1176
Db      1021 ACTTTGGACCGGATTTCAAGCTTCAATCAGCCCTTTCTAACATGACCAACCAAGACACTA 1080
Qy      1177 atgagctacgtgagagagatcaaaacagcagctgtttgagaacctagaaatgctgcgcagc 1236
Db      1081 ACGACTACTCGAGAAATCAAGACAGCGCTCTTTGAGAACTTGAGAGATGCTGCCCATG 1140
Qy      1237 cactgtgggtccaaacagcagcagctcctgagagacgcacccctgagagagatgagatg 1296
Db      1141 CCCCCTGGGGTCCAGATGACAGGCGCATCTGAGAGCGCATCCCGAAGAGAGTGGGAGATG 1200
Qy      1297 aggaagaaagacagccctgacagcagcagcagctgctgctctgacaaacgaattgct 1356
Db      1201 AAGATGAGAGAGACCCCTGCAAAACGATCTCCATCTGCTCTGTATTAACGCAATTGCT 1260
Qy      1357 gtcgaagaaagcttctcagctatctgagagagagagagagagagagagagagagagagag 1416
Db      1261 GTGAGAGAGAGTTCGCGACCTCAGATGAGAGAGAGAGAGAGTGTGCAAGAACTTTCTA 1320
Qy      1417 acttaaaaaaacagagagagatcaaaacagagatgataaagaagaagaagagagagagag 1476
Db      1321 ACTTTAAAAAACCCAAAGAGTTAAACAGAGATGAGAAAGAGAAAGAGAGAGAGAGAGAG 1380
Qy      1477 agaaagagatccagcagagagagagaaacaaagagagagagagagagagagagagagag 1536
Db      1381 AAAAAGAAAGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Qy      1537 aggaagagagcagagctgagctgagagagagagagagagagagagagagagagagag 1596
Db      1441 AAGAAGAGGTCAAGTTGGCGCTG-ACCAAGAGTCTGCAAGGCCCATCTTCTCCCAAGTTCT 1499
Qy      1597 cagcttctt 1606
Db      1500 CACTCTCCT 1509

```

RESULT 10
AP000450
LOCUS 81975 bp DNA linear HTG 30-MAY-2000
DEFINITION Homo sapiens chromosome 11 clone XXP1-307 map 11q23, WORKING DRAFT
SEQUENCE, 5 unordered pieces.
ACCESSION AP000450
VERSION AP000450.3 GI:8118763
KEYWORDS HTG; HTGS; PHASE1; HTGS; DRAFT.
SOURCE Homo sapiens DNA, clone:XXP1-307.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 81975)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 81,975 genomic DNA of 11q23
Published Only in Database (1999) In press
2. (bases 1 to 81975)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (13-SEP-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
Kitsato Univ., 1-15-1 Kitsato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail:hattori@gsc.riken.go.jp,

COMMENT

```

URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)
On May 31, 2000 this sequence version replaced gi:7630227.
-----
Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
-----
Project Information
Center project name: Humdrift11
Center clone name: XXP1-307
-----
Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 79702 bases at least Q40
Consensus quality: 80980 bases at least Q30
Consensus quality: 81428 bases at least Q20
Insert size: 81575; sum-of-contigs
Quality coverage: 5.78x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
5 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs 'N', but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1 35184 contig of 35184 bp in length
35285 51779 contig of 16495 bp in length
51880 63010 contig of 11131 bp in length
63111 73886 contig of 10776 bp in length
73987 81975 contig of 7989 bp in length
Sequence updated (18-Apr-2000)
NOTE: This is a 'working draft' sequence. It currently
consists of 5 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of 'N', but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 35184: contig of 35184 bp in length
35185 35284: gap of 100 bp
35285 51779: contig of 16495 bp in length
51780 51879: gap of 100 bp
51880 63010: contig of 11131 bp in length
63011 63110: gap of 100 bp
63111 73886: contig of 10776 bp in length
73887 73986: gap of 100 bp
73887 73986: gap of 100 bp
73987 81975: contig of 7989 bp in length.
Location/Qualifiers
1. 81975
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q23"
/clone="XXP1-307"
1. 35184
/note="assembly-fragment"
35285. 51779
/note="assembly-fragment"
51880. 63010
/note="assembly-fragment"
63111. 73886
/note="assembly-fragment"
73987. 81975
/note="assembly-fragment"
BASE COUNT 24973 a 15476 c 15771 g 25355 t 400 others
ORIGIN

```


REFERENCE 6 (bases 1 to 146841)
 AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (09-MAY-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On Aug 11, 2000 this sequence version replaced gi:7630875.

----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: saplens@wustl.wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0172C16

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc.

SOURCE INFORMATION:
 The RPc1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P. Y., Zhao, B., Frengen, E., Tateo, M., Catanesse, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (http://bacpac.med.buffalo.edu) VECTOR: pBac3.6

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is RP11-720G3; the clone sequenced to the right is RP11-643E23. Actual start of this clone is at base position 1 of RP11-172C16; actual end is at base position 146841 of RP11-172C16.

The region from base position 130951 to 131438 represents the best possible assembly of an imperfect TA repeat. Restriction digest information from EcoRV, HindIII, and BamHI indicates that approximately 150 to 250 bases may be missing from the final sequence.

FEATURES

Source

1. 146841
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="RP11-172C16"
 /clone_lib="RPc1-11"
 418..575
 /rpt_family="MER1_type"
 774..946
 /rpt_family="L1"
 1114..1313
 /rpt_family="L1"
 1365..1452
 /rpt_family="ERV"
 1425..1510
 /rpt_family="ERV"

repeat_region	1541..2262	/rpt_family="L1"	EST AW009077 (NID:95857855) ws76g06.x1"
repeat_region	2269..2957	/rpt_family="L1"	EST AM844097 (NID:97938080)"
repeat_region	8426..8655	/rpt_family="L1"	EST AM841802 (NID:97935785)"
repeat_region	8656..8957	/rpt_family="L1"	
repeat_region	8958..8996	/rpt_family="Alu"	
repeat_region	9029..9453	/rpt_family="L1"	
repeat_region	10059..10457	/rpt_family="L1"	
repeat_region	10468..10610	/rpt_family="MALR"	
repeat_region	10659..10807	/rpt_family="L1"	
repeat_region	10843..11005	/rpt_family="L2"	
repeat_region	12544..12725	/rpt_family="ERV1"	
repeat_region	12830..13126	/rpt_family="L1"	
repeat_region	13433..13674	/rpt_family="Alu"	
repeat_region	14419..14587	/rpt_family="L1"	
repeat_region	15199..15545	/rpt_family="MIR"	
repeat_region	16012..16077	/rpt_family="L2"	
repeat_region	17372..17741	/rpt_family="L2"	
misc_feature	17555..18180	/note="similar to 17555..18180"	EST AW009077 (NID:95857855) ws76g06.x1"
misc_feature	17696..18060	/note="similar to 17696..18060"	EST AM844097 (NID:97938080)"
repeat_region	17999..18157	/note="similar to 17999..18157"	EST AM841802 (NID:97935785)"
repeat_region	18220..18324	/rpt_family="L2"	
repeat_region	18325..18401	/rpt_family="L1"	
repeat_region	18408..18538	/rpt_family="Alu"	
repeat_region	18539..19181	/rpt_family="Alu"	
repeat_region	19454..19661	/rpt_family="L1"	
repeat_region	19705..20140	/rpt_family="L1"	
repeat_region	20162..20624	/rpt_family="L1"	
misc_feature	20371..20690	/note="similar to 20371..20690"	EST AI347274 (NID:94084480) tc05d02.x1"
repeat_region	22353..22774	/note="similar to 22353..22774"	EST AM844024 (NID:97938007)"
repeat_region	23334..23715	/rpt_family="MALR"	
repeat_region	23738..23802	/rpt_family="ERV1"	
repeat_region	23803..24380	/rpt_family="MALR"	
repeat_region	24381..24774	/rpt_family="ERV1"	
repeat_region	24925..25100	/rpt_family="MALR"	
repeat_region	25169..25250	/rpt_family="L2"	
repeat_region	25524..25566	/rpt_family="MER53"	
repeat_region	25567..25668	/rpt_family="L2"	

Db	63751	TCCTCAGAGGGGCTACCTAGATGTCCTTAATCTGATCATCAAAAGGCATGGCCAGTGTGTGG	63692
Qy	967	aattgtcaagaagcgtttaacctctgacctatgctgtagtctgtaggaagcggltgtaacacatc	1026
Db	63691	AATCTGTCAAGAGTGTCAACACCTGCTATGTTATTTGATGTGCACAAAGCTGTTCACCATTTT	63632
Qy	1027	gtcaagctgcccgggtg--ctgagacatataagaacagctgtgagccctggtatgaacgagatccc	1084
Db	63631	ATTAATGTTGCTTAGTGGCCCTGGAAATATAGAAAGAGCTGTGGCTCTGTGGATATATGAGATCCC	63572
Qy	1085	taatgagcttcatacaatactacttltgaatctctgtgacacagatcttgaacotcacat	1144
Db	63571	TATGAGCCTTCATACATGATTTACTTTGAATTTGACCATATTTTCAACCCCTACAT	63512
Qy	1145	cagtccttccaatatagaataccagaacaagaatgaactcctgtagaagaatcaacaagcy	1204
Db	63511	CAGTCCCTTCCATATGACTATGACCTAACCAACACTATGATGCTGTGGAGAAATTAACAGTGG	63452
Qy	1205	actgttttagaaccttagaattgctgcgcgacagactgggtccaaacgcaagcgattcc	1264
Db	63451	GCTGTTTGAAACCTATGATATCCAGGCCCTACCTGCAGTGCACCAATGACGATTA-T-C	63393
Qy	1265	tgaagacgcacatcccttgagagagatgtagatgaagaacgaacacccctgcacaagcgat	1324
Db	63392	TGAGATGCTGTGTCCTTCAAGAGAGTGGCGATAGAAATGAAGAAGACCTGCAGACACTCAT	63333
Qy	1325	ctgcattctgctctctcttgacaacaagaattgctctgtgaagaagaagttctccgattctgaaga	1384
Db	63332	CTCCATCCGCTTCTCTGACAGATGAAATTAACCTGTAGAGAAAGATTTCTTGACTGTGATGA	63273
Qy	1385	ggaagagagagagggggccgaagaactcttccacttcaaaagaacgaagagatgaac	1444
Db	63272	GGAGGAGAGCTGCTGTGACAGATCTCTCCATCTTCAAAAAGCCAAAGGATTTAAAC	63213
Qy	1445	agaagatgaaaaagaagaagaccagagagagaagaagaatccacgaagagagaagaac	1504
Db	63212	AG----GATGAAAAAAGAGAGACTCAGAGATGAGAAATGTCTACACAAAGAGGAGA-AAAAT	63158
Qy	1505	caaggaagagaagaagccagaagcccaaaagggtcaagaagaagagccaaagtggcctgaatga	1564
Db	63157	GAAAGAGAGAGAAAGCCAGAGAGCCAAAAGAGGTCAAGAGAGAGGTCAAGCTGCCCTGAGGAA	63098
Qy	1565	ctctcccaagctctgctctcctgctgtagtgcctcccaagttcttctccc	1610
Db	63097	CCTCCCTGCTGTGCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTTCTCTCC	63052
RESULT	13		
AC105705/c			
LOCUS			
DEFINITION	AC105705	252957 bp	DNA
SEQUENCE	Rattus norvegicus chromosome Idm4 clone CH230-67A16, WORKING DRAFT		
ACCESSION	AC105705		
VERSION	AC105705.1	GI:18092928	
KEYWORDS	HTG: HTGS_PHASE1; HTGS_DRAFT.		
SOURCE			
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
REFERENCE	1 (bases 1 to 252957)		
AUTHORS	Muzny,D.M., Adams,C., Adlo-Oduola,B., Ali-osman,F.R., Allen,C.,		
	Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Banks,T., Barbarta,J.,		
	Benton,J., Bismage,K., Blankenburg,K., Bonnin,D., Bouck,J.,		
	Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,		
	Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,		
	Carter,M., Cayazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,		
	Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C.,		
	Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,		
	DeVey-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,		
	Duan,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,		
	Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,		
	Elhaj,C., Escoto,M., Falls,T., Ferraruto,D., Flagg,N., Ford,T.,		

Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
 Garza, N., Gill, R., Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S.,
 Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,
 Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Hollway, C.,
 Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,
 Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
 Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
 Kovar, C., Kirov, C., Kuresh, A., Landry, N., Leal, B., Lewis, L. C.,
 Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
 Loussaged, H., Lozado, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
 Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
 Martinez, E., Massey, E., Mawhinney, E., Mcleod, M. P., Meador, M.,
 Mel, G., Metzner, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
 Morgan, M., Morris, S., Mose, M., Neal, D., Newton, J., Newton, N.,
 Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokoko, S.,
 Ogih, M., Okunoye, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,
 Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L.,
 Qulles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Roife, M.,
 Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshbar, N.,
 Slisom, I., Sodergren, E., Sonake, T., Sparks, A., Stanley, H.,
 Stone, H., Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K.,
 Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
 Thomas, S., Usmani, K., Vasquez, L., Vera, Y., Villalton, D., Vinson, R.,
 Wall, R., Wang, S., Ward, Moore, S., Warren, R., Wallington, C.,
 Wallington, S., Williams, G., Williamson, A., Wlezyk, R., Wooden, S.,
 Wolter, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 252957)
 Worley, K. C.
 Direct Submission
 Submitted (09-JAN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GNPA
 Center clone name: CH230-67A16
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329first call to
 findPhrapList
 Consensus quality: 208655 bases at least Q40
 Consensus quality: 216460 bases at least Q30
 Consensus quality: 222170 bases at least Q20
 Estimated insert size: 217567; sum-of-coverage estimation
 Quality coverage: 0x in Q20 bases; agarose-gel estimation
 Quality coverage: 3x in Q20 bases; sum-of-coverage estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 70 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 8561: contig of 8561 bp in length
 * 8562 8661: gap of unknown length
 * 8662 17870: contig of 9209 bp in length
 * 17871 17970: gap of unknown length
 * 17971 28613: contig of 10643 bp in length
 * 28614 28713: gap of unknown length
 * 28714 36948: contig of 8235 bp in length
 * 36949 37048: gap of unknown length
 * 37049 44586: contig of 7538 bp in length
 * 44587 44586: gap of unknown length

[illegible]

VERSION	AF04169.1	GI:2829213
KEYWORDS	chicken.	
SOURCE	Gallus gallus	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Aukosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus	
REFERENCE	1 (bases 1 to 1646)	
AUTHORS	Sun,J.M., Chen,H.Y., Moniva,M., Samuel,S. and Davie,J.R.	
TITLE	Purification and characterization of chicken erythrocyte histone deacetylase 1	
JOURNAL	Biochemistry 38 (18), 5939-5947 (1999)	
MEDLINE	99249807	
PUBMED	10231548	
REFERENCE	2 (bases 1 to 1646)	
AUTHORS	Sun,J.M., Chen,H.Y. and Davie,J.R.	
TITLE	Direct Submission	
JOURNAL	Submitted (22-JAN-1998) Biochemistry, University of Manitoba, 770 Bannatyne Avenue, Winnipeg, MB R3E 0W3, Canada	
FEATURES	location/Qualifiers	
source	1..1646	
	/organism="Gallus gallus"	
	/db_xref="taxon:9031"	
	/cell_type="definitive erythrocyte"	
	/dev_stage="15 day embryo"	
	/note="also in adult liver"	
gene	1..1646	
	/gene="HDAC1"	
	48..1490	
	/gene="HDAC1"	
	/note="associated with nuclear matrices of primitive and definitive erythrocytes"	
	/codon_start=1	
	/product="erythrocyte histone deacetylase"	
	/protein_id="AAC0504.1"	
	/db_xref="GI:2829214"	
CDS	/translation="MALTCGRKRVCCYIYGGDNGNYIYGGGHPKPRIRMTNHLIN YGLRKMEIRYPRKANAEMTKYHSDDYIFLRSIRPDNMSSEYSKQMRNVSEDCPV FDLGFEFCOVSAGSIVASAVKLAKQODIDAVNMAAGLHAHKBSASGFCYNDIVLAI LELIKLHQRYLVIDIDIHGGVGEAFYTDRTVMTSPFKRYGEYFETGDLRIGAGK GKVAAYNVPLRDGIDDESVEAIFKPVISKWNETPOPFAVVLQSGDSLSGDRIGCFNL TIKHAKCVFVKSFENLPMILGGGGITINVARCQVETPAVALDPIELPELYNDFE FYRPDPKRLHSPSNTNONTNMYLEKIKORLEPMLPHAGYQOMPIPEAYQDS SGDEEDEDPEKRISIRNSDKRISODEEFSSEDEGEGKRNVAFAKAKRYKTEERE EEEKRDEKEERKAEEKAEPKGVKEEKS"	
BASE COUNT	461 a 346 c 452 g 387 t	
ORIGIN		
Query Match	61.9%	Score 997.4; DB:5; Length 1646;
Best Local Similarity	80.2%	Pred. 5.2e-260;
Matches 1185; Conservative	0; Mismatches 286; Indels 6; Gaps 1;	
Db	83 gggtagaccacgcgaggaggcgagcaagaatgcgacagcgacgagcgaggaagt 142	
	20 ggggtctggggcgtggggagagcgagaaatggcctgacgcaggggacgaagcgaaagt 79	
QY	143 ctgttactactacgacgaggagtggtgaattactattatgagcaagcgacccaatgaa 202	
Db	80 ctgctactactacgacgagcggtgatgttggaactatattactatgcccagggcgaccccatgaa 139	
QY	203 gcttaccgaatcgcgactgacatcaatttctgttcaacatgcttcacgcgaaaaat 262	
Db	140 gccccacagatccggatgacccacacacttactgctgactacagccttttaccaggaagt 199	
QY	263 ggaatctatcgcgccctcccaagccaatgctgtgagagatgacccaagtaccacaagatga 322	
Db	200 ggagatattatgccctccacaaaggcgagcgagagatgacccaagtaccacagatga 259	
QY	323 ctacattaatcttgcgctcaatcgtccagatatacatgctcgagatcagaagaagcagat 382	
Db	260 ctacatcaaatcttcgagatccatccgccgacgaacacatgctctgagtacacacgaagat 319	

383 gcagaagattcaacgttgtagagactgcacgattcgatgctgtttagtctgtca 442
320 GCAAGATTTTAACTCGGAGGAGACTGCCGTGTCATGGGCTTGGATTCTGTCA 379
443 gcttctactgtgtgtctgtgcaaatgtgttaacttaataaagcagagagat 502
380 GCTCTCTGCTGGAGGCTCCGTTGCCAGCGCTGTGAACTGAAACAAGACAGATAT 439
503 cgcgtgtaattggcctggggagcctgcacacatgcagaagatccgagatctgtctc 562
440 TGCTGTGAATTTGGCAGAGAGCCTTACCACAGCTTAAGAGTGGAGGCTTGGCTTGTG 499
563 ttaagtcataatatactgtctgtgcaatccctgaactgttaagatataccagaagtgct 622
500 TTATGTCAACATATATGCTTCCTGCTATCTTGAAGTCTTAAAGTATCACAGAGGCTCT 559
623 gtacattacatataatataccatctgtgagcggctggaaagagcctctcaacagaga 682
560 GTATATTCACATTGTATATTCACATGAGATGCTGTGAGAGAGAGCCTTCTATACCACAGA 619
683 ccgggtacatgactgtcttctcttcaatgaatgagagactcccaagaaactggagacct 742
620 CCGTGTATGACCGCTGCTCTTCAATAGATGAGAGTACTTCCAGAGAACAGGGGACCT 679
743 acggagatccgggctgagcaagaagatataatgctgttaactaccgctccgaagcgg 802
680 GCGGACATATGCTGACAGCAAGCAAACTACTGCTCACTATCCCTCCGGAGATGG 739
803 gattgatacagatcctctctgagccatttcaagccggtctgtcaagaatgagat 862
740 GATTGATATGATGCTCTCAAGAGCAATATCAAGCGGTGATATTAAGATGATGAGAC 799
863 gtccacagctagtcggtgtcttcaagtgctcgaactcccatctcgtggagacgtt 922
800 ATTCACAGCTAGTGTGAGTGTCTGCTCACTGGCGGTGAGATTCGTCCGGGAGACAGCT 859
923 aggtgtctcaatctatctcaagaagcagcgaagtgtgtgaaattgtcaagacct 982
860 GGGTGTGTTTATCTGACCATCAAAAGTCAAGCTGAGGTGTGTCATAAAGTGT 919
983 taacctgtctatgtcgtgagcggtgtgtgtacacattgtcaacgttgcacggtg 1042
920 TAAATTGCTTATGCTGATGCTGGAGAGGAGTGCTATACATCCCAACGTTGGCCAGATG 979
1043 ctggaacataatagacagctgtgcccgtgatacggagataccctaaatgagctcaca 1102
980 CTGACACTATGAGAGCTGCTGGGCTTGGACACTGAGATCCCAATGAGCTCCCATATA 1039
1103 tgactacttgaatacttgaacacagatltcaagctccacatcaatctcctcaatagac 1162
1040 TGACTATTTTATGATCTTGGACCAAGCTTAAAGTGCACATCACTCCCTCAACATGAC 1099
1163 taacagaacacagatgagtagtacccttgagagaatacaaacagcagctgttgaacattag 1222
1100 CAACCAAGATACCAATGAGTATCTTCGAGAAAGTCAAGACGCTCTCTTGGAGAACTGCG 1159
1223 aatgtgcgcagcagcacttgggtgtccaaacgcagcgatltcctgagagacgcacatccctga 1282
1160 CATGCTGCTCATGCTGCTGGGCTCCAGATGACCAATTCCTGAGGATGCTGTCAAGA 1219
1283 gtagagtgtagatgagagcagagacccctgacacgcgcatctgactctctctga 1342
1220 AGACAGTGGGATGAGAGAGAGAGATCTGAGAAAGCGCATTTCAATCCGCAATTTGCA 1279
1343 caaagaacatgcctgtgagagagatltctcgatctgaagagagggagagggggccg 1402
1280 TAAAGGAATATCTCTGTGTAAGAAATTTCTGTACTCTGAAGTGAAGGGAGAGGAGCG 1339
1403 caagaactcttcaactcaaaaaaagcagaagatltcaaaaacagagatgtgaaaaagaa 1462
1340 CAAAAATTTGTCACACTTTAAGAAAGGCCAAGGCTGTGAACACAGAGAGAAAGAGAGA 1399
1463 agaccacagagagaaagaaagatataccgaagagagagaaaccaaagagagagaaagccaga 1522

Db 1400 -----GGAGGAGAGAAAGATGAGAGAAAGAGAGAAAAAGAGAGAAAGCCGA 1453
Oy 1523 agccaagggatcaaggagagagccaaagtggccgga 1559
Db 1454 ACCCAAGGGGGTGAAGAGAGAGACAAATCCACCTTA 1490
RESULT 15
AF039751
LOCUS AF039751 1678 bp mRNA linear VRT 18-AUG-1999
DEFINITION Gallus gallus histone deacetylase-1 mRNA, complete cds.
ACCESSION AF039751
VERSION AF039751.1 GI:2791683
KEYWORDS
SOURCE
ORGANISM Gallus gallus
Chickens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
1 (bases 1 to 1678)
Takami,Y., Kikuchi,H. and Nakayama,T.
Chicken histone deacetylase-2 controls the amount of the Igm
H-chain at the steps of both transcription of its gene and
alternative processing of its pre-mRNA in the DF40 cell line
J. Biol. Chem. 274 (34), 23977-23990 (1999)
JOURNAL
MEDLINE 99377029
REFERENCE
AUTHORS 2 (bases 1 to 1678)
TAKAMI,Y.
TITLE Direct Submission
Submitted (23-DEC-1997) Biochemistry, Miyazaki Medical College,
Kiyotake Kihara 5200, Miyazaki 889-16, Japan
location/Qualifiers
1. 1678
/organism="Gallus gallus"
/db_xref="taxon:9031"
/cell_line="DF40"
58. 1500
/note="HD-1"
/product="histone deacetylase-1"
/protein_id="AAB96923.1"
/db_xref="GI:2791684"
/translation="MALTGQTKRKYCYIIDGVDGNYTYGGCHPKPHIRIRNTNLLN
YGLRKMETIRPKANAEEMTKYHSDVIFKRLSIRPNMSEYSKQOFNVNGDCEPV
FDGLFEFQISAGGSVAAYKLNKOQDIAVNMAGLPHAKSSASGFCYVNDIVLAI
LELKHORVLYIDIDIHGDSVEAEFTDRTVMVSPHXYGEFPGDRLDGAGK
GKYAVNVPILRDGIDDESVEAIFKPVLSKWEFTOPSAVVVIGSDSISGDRICFNL
TIKHKACVEYKSFNPLMLGGGGYIRRVAVCWYETVALDETIRNPLRYNDYF
EYFGDFELHISPSNMTNONTNLETKIKORLENLRLPHAPGVOMOPTEDAVQED
SGDEEEDPEKRIISIRNSDKRISDDEERSDSEDEGEGRRKNVANFKAKRVTEERE
EEKKDEKEEERKAEPKGVETEEST"
BASE COUNT 462 a 357 c 466 g 393 t
ORIGIN
Query Match 61.9%; Score 997.4; DB 5; Length 1678;
Best Local Similarity 80.2%; Pred. No. 5.2e-260;
Matches 1185; Conservative 0; Mismatches 286; Indels 6; Gaps 1;

Db 210 GGAGATATATCCCTCCACAAAGCGGAAACCGAGAGATGACCAATACACAGTATGA 269
QY 323 ctacatataattcttgcgtccatccgtccagataaactgctggagttacagaagaagt 382
Db 270 CTACATCAAAATTCCTGAGATCCATCCGCCAGCAACATGTCTGATACAGCAAGAGAT 329
QY 383 agagagatcaacgcttggagagactgtccagtatccatggcctgttgaagttcgtca 442
Db 330 GCAAAGATTTAAAGTGGGGAGAGACGCGCTGTTCATGGGCTGTTCAGTTCGTGA 389
QY 443 gttgttactggtgttctgtgtgcaagtgctgtgaaacttaataagcagcagaagacat 502
Db 390 GCTCTCTGCTGGAGGCTCCGTTGCCAGCCTGTGAAGCTGAACCAAGCAACAGACATAT 449
QY 503 cggcgtgaattggcgtgggggctgtccatccatcaagaagtgccagggcctgtgtctg 562
Db 450 TCGCTGAATTTGGGCGAGGAGCTTCACACGCTAAAGAGTGAGGAGCTTCTGGCTTG 509
QY 563 ttaactcaatgatalatcgtctgtgccaaccctggaaactgtcaaaagtatccacagaagtgct 622
Db 510 TTATGTCAACGATATATGCTGCTGCTATCTTGAGAGCTTAAAGTATCACAGAGGCTCT 569
QY 623 gtaacatgacatgtgataatcaacatggtgagcggtggaagagcgcttccacacagga 682
Db 570 GTATATTTGACATTGATTTACCATGAGATGAGATGAGTGTGAGGAGAGCCTTATACACAGA 629
QY 683 ccgggtgcaactgtgttcccttccatgaatgaaagtactccacagagacttggagacct 742
Db 630 CCGTGTGATGACCGTCTCTTTCATTAAGTATGAGAGTACTTCCAGGAACAGAGGACCT 689
QY 743 acggagatccggggctgtgcaaaagaaagatataatgctgttcaactaccgcctccagaacgg 802
Db 690 GCGGGCATTTGGTGCAGGCAAAAGCAAAATATATGCTGTCAACTATCCCTCCGGGATGG 749
QY 803 gatitgagcaggtctcctagagagccattcaagccggtcagtgccaaagtatgagat 862
Db 750 GATTGATGATGATCTCTACAGAGGCAATATTCAGCCGGTATATCTAAAGTGTGAGAGAC 809
QY 863 gtccagacctagtgcggtgtcttcaagtggtcagactccctacccctacccgtggagatcgct 922
Db 810 ATTCCAGCTATGAGACGTGTGTCTGTGCAAGTGGGCTGGATTCCTCTCCGGGAGACGCT 869
QY 923 aggttgctcaactctatctatcaaaagagacgccaagtgtgtgaaattgtcaagaagctt 982
Db 870 GGGTTGTTTAAATCTGACCATTAAGGTCATCCAAAGTGTGAGTGTGCAAAAGTTT 929
QY 983 taacgtcctatgctgatatctgtggagcggtgttaaccattcgtaaagttgcccgtg 1042
Db 930 TAAATTTGCTTATGCTGATCTGGAGAGAGTGGCTATATACGATCCGCAACGTGGCCAGATG 989
QY 1043 ctggacataatgagacagctgtgtgcccctgtgatacagagatccctaaatgagcttccatacaa 1102
Db 990 CTGGACCTATGAGACGTGTGTGCTTGTGACACTGAGATCCCAATGAGCTCCCATATTA 1049
QY 1103 tgaactacttgaactcttggaccagatttcaagctccacatcagtccttccaataatgac 1162
Db 1050 TCACATATTTGACTTGTGGACCAAGACTTAAAGCTGCACATCAGTCCCTCAAAACATGAC 1109
QY 1163 taacagaaacagaaatgagttaccttgagagagatcaaacagagactgtttgagaacctag 1222
Db 1110 CAACCAAGAAATACCAAGATATCTCGAAGAAATCAAGCAGCGTCTTTGAGAAATCTGCG 1169
QY 1223 aatgtcgccagcagcagcctggggttccaaacgagagacttccgtgagagagccatccctga 1282
Db 1170 CATGCTGCTCATGCGCTGCGGCTCCAGATGCAAGCCAAATTCCTGAGAGATGCTGTTAGGA 1229
QY 1283 ggaagatggcgaatgagagcagaacgacccctgaacaagcgatcgcgtcgtcctga 1342
Db 1230 AGACAGTGGGAGTGAAGAAAGAAAGATCTCTGAAAGCGCATTTCAATCCGCAATTTCTGA 1289
QY 1343 caaacgaattgcctgtgaggaagagttcgcgattcgaagagagagagaggggccc 1402

Db 1290 TAAAGAAATATCTGTGATGAAGAAATTCCTGACTCTGAAGATGAAGGGAGAGGCGC 1349
QY 1403 caagaactcttccacttcaaaaaaagccagaagagttcaaaacagagagatgaanaagaa 1462
Db 1350 CAAAAATGTTGCCAACTTTAAGAGAGCCAAAGCTGTGAAGAACAGAGAGAGAGAGAG 1409
QY 1463 agaccagaagagaaagaaatcacccgaagagagagaaacaaagagagagagagccaga 1522
Db 1410 -----GAGAGAGAGAGAGATGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1463
QY 1523 agccaaaagggtcagaag 1559
Db 1464 ACCCAAGGGGTGAAG 1500

Search completed: June 8, 2002, 12:37:02
Job time: 10915 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2002, 09:41:03 ; Search time 279.1 Seconds
(without alignments)
9910.244 Million cell updates/sec

Title: US-09-817-538-2
Perfect score: 1611
Sequence: 1 atgtctggggctcttcgccg.....tccctcagtttcttccccc 1611

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 segs, 858457221 residues
al number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 11
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq_032802:*

1: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS5/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1596.6	99.1	1611	22 AAC89554	Human histone deac
2	1499.6	93.1	2163	21 AAF16188	Human prostate can
3	1498	93.0	2111	17 AAT12940	RBDL transcription
4	1434.6	89.1	1449	18 AAT86371	cDNA encoding a hi
5	1378	85.3	1582	23 AAS81302	DNA encoding novel
6	818	50.8	1997	24 AB199512	Mouse ischaemic co
7	805.2	50.0	1985	20 AA232066	Human MEFH2 relate
8	805.2	50.0	1985	22 AAC90323	U1814 cDNA clone.
9	805.2	50.0	1985	22 AAC89555	Human histone deac

10	729	45.3	2145	23 ABL01897	Drosophila melanog
11	640.6	39.8	4755	23 ABL01896	Drosophila melanog
12	502.6	31.2	1943	20 AAX90840	Maize histone deac
13	501.8	31.1	2019	20 AAX90839	Maize histone deac
14	499.6	31.0	1990	21 AAZ58259	Rice histone deace
15	498.2	30.9	1807	22 AAF80350	Nucleotide sequenc
16	480.2	29.8	1954	22 AAC89556	Human histone deac
17	477	29.6	1805	21 AAZ58260	Soybean histone de
18	460.8	28.6	1576	20 AAX90841	Maize histone deac
19	459	28.5	1826	20 AAX90837	Maize histone deac
20	434	26.9	1662	23 ABL04121	Drosophila melanog
21	433.4	26.9	1609	21 AAC46043	Arabidopsis thalia
22	433.4	26.9	1611	21 AAC33022	Arabidopsis thalia
23	431.8	26.8	1800	22 AAF80351	Nucleotide sequenc
24	413	25.6	586	22 AAS25097	Human ovarian PCR-
25	413	25.6	586	22 AAH83742	Human ovarian tumo
26	402	25.0	576	22 AAS25098	Human ovarian PCR-
27	402	25.0	576	22 AAH83743	Human ovarian tumo
28	392.4	24.4	600	21 AAL16397	Human colon cancer
29	386.8	24.0	402	22 AAH55632	Human breast tumou
30	369	22.9	3793	23 ABL04120	Drosophila melanog
31	329.2	20.4	1475	20 AAX90838	Maize histone deac
32	318.2	19.8	2009	22 AAD17396	Human bone marrow
33	286	17.8	286	23 AAS57741	cDNA #417 encoding
34	284	17.6	3826	23 ABL01928	Drosophila melanog
35	235.8	14.6	241	22 AAS47345	Human breast cancer
36	235.8	14.6	241	22 AAF17915	Human breast cancer
37	201.6	12.5	1654	21 AAZ93329	Human histone deac
38	201.6	12.5	2046	22 AAI59530	Human polynucleoti
39	200	12.4	1367	22 AAH75178	Nucleotide sequenc
40	200	12.4	1682	22 AAH25122	Nucleotide sequenc
41	195	12.1	379	18 AAT86372	Partial cDNA encod
42	169.4	10.5	375	18 AAT86373	Partial cDNA encod
43	158.4	9.8	511	22 AAT86374	Human cervical can
44	151.8	9.4	227	18 AAT86374	Partial cDNA encod
45	137.4	8.5	470	22 AAH71919	Human cervical can

ALIGNMENTS

RESULT 1	
AAC89554	03-MAY-1999; 9909-0312287.
ID AAC89554 standard; DNA; 1611 BP.	
AC AAC89554;	
XX	
DT 08-MAR-2001 (first entry)	
XX	
DE Human histone deacetylase HDAC-1 coding sequence.	
XX	
KW Histone deacetylase; HDAC-1; HDAC-2; HDAC-3; HDAC-4; HDAC-5; HDAC-C;	
KW HDAC-D; cell cycle; tumorigenesis; cancer; inhibitor; antisense;	
KW gene therapy; ds.	
XX	
OS Homo sapiens.	
XX	
PN WO200071703-A2.	
XX	
PD 30-NOV-2000.	
XX	
PF 03-MAY-2000; 2000WO-IP01252.	
XX	
PR 03-MAY-1999; 9909-0312287.	
XX	
PA (METH-) METHYLGENE INC.	
XX	
PI Macleod AR, Li Z, Besterman JM.	
XX	
DR WPI; 2001-016407/02.	
XX	
DR P-PSDB; AAB49954.	
XX	
PT Antisense oligonucleotide that inhibits expression of a histone	


```
Db 721 ttcaagccggtatgtccaagtaatgagatgtccagccctagtcggtgtcttaag 780
Qy 891 tttgtctgaatccctctctctctctctctctctctctctctctctctctct 950
Db 781 tttgtctgaatccctctctctctctctctctctctctctctctctctctct 840
Qy 951 caagcgaagtgtgtgaattgtctcaagagcttaaccctctctctctctctctct 1010
Db 841 caagcgaagtgtgtgaattgtctcaagagcttaaccctctctctctctctctct 900
Qy 1011 gttgtttacaccatctgtaagctgtgcccgtgtctggaatgaaacagctgtg 1070
Db 901 gttgtttacaccatctgtaagctgtgcccgtgtctggaatgaaacagctgtg 960
Qy 1071 gatacggagatcccttaagcttgaatctatatacagcttgaatcttgaacag 1130
Db 961 gatacggagatcccttaagcttgaatctatatacagcttgaatcttgaacag 1020
Qy 1131 ttcaagctccacatcagctctctctcaatactgactaaccagaacacgaatg 1190
Db 1021 ttcaagctccacatcagctctctctcaatactgactaaccagaacacgaatg 1080
Qy 1191 aagatcaaacagcagctgtttgagaaccttagaatgtctgcgcagcagctgtg 1250
Db 1081 aagatcaaacagcagctgtttgagaaccttagaatgtctgcgcagcagctgtg 1140
Qy 1251 acgcagcagatctcgtgagagagccatccctctgagagagtgagatgagag 1310
Db 1141 atgcagcagatctcgtgagagagccatccctctgagagagtgagatgagag 1200
Qy 1311 cctgcaagcagcagctctctctctctctctctctctctctctctctctct 1370
Db 1201 cctgcaagcagcagctctctctctctctctctctctctctctctctctct 1260
Qy 1371 tccgattctgaagagagagagagagagagagagagagagagagagagagag 1430
Db 1261 tccgattctgaagagagagagagagagagagagagagagagagagagagag 1320
Qy 1431 aagaagatcaaacagcagctgtgaaagaaagaaagaaagaaagaaagaaag 1490
Db 1321 aagaagatcaaacagcagctgtgaaagaaagaaagaaagaaagaaagaaag 1380
Qy 1491 gaagagagagagagagagagagagagagagagagagagagagagagagag 1550
Db 1381 gaagagagagagagagagagagagagagagagagagagagagagagagag 1440
Qy 1551 ttggcctga 1559
Db 1441 ttggcctga 1449

RESULT 5
AAS81302
ID AAS81302 standard; cDNA: 1582 BP.
AC AAS81302;
DT 13-FEB-2002 (first entry)
DE DNA encoding novel human diagnostic protein #17106.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX OS
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX
```

```
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
DR WPI: 2001-639362/73.
P-PSDB: ABG17115.
PT New isolated polynucleotide and encoded polypeptides, useful in
PR diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PR biodiversity
PS Claim 1; SEQ ID NO 17106; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
CC
SQ Sequence 1582 BP; 422 A; 376 C; 441 G; 343 T; 0 other;

Query Match 85.5%; Score 1378; DB 23; Length 1582;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 1475; Conservative 0; Mismatches 40; Indels 13; Gaps 6;

Qy 97 ggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 156
Db 55 ggaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 114
Qy 157 acggagatgttgaattactattatgtgac-aaggccacccaatgaagcttc-accgaa 214
Db 115 acggagatgttgaattactattatgtgagcagcagcagcagcagcagcagcagcag 174
Qy 215 ccgcatgactataatttgcgtctcaactatggtctctcaacgaaataatggaattatcg 274
Db 175 ccgcatgactataatttgcgtctcaactatggtctctcaacgaaataatggaattatcg 234
Qy 275 cccctacaagcacaatgtctgagagagatgaccagaatgaccagcagcagcagcagcagc 334
Db 235 cccctacaagcacaatgtctgagagatgaccagaatgaccagaatgaccagaatgacc 294
Qy 335 ctggcgtccatccctccagataactgttgcggagataccagcagaatgacatgacaga 389
Db 295 ctggcgtccatccctccagataactgttgcggagataccagcagaatgacatgacaga 354
Qy 390 ttcaacgttggtagagactgtcc-agtattcgatggcgtg-tttagttctgtcagtg 446
Db 355 ttcaacgttggtagagactgtccagatgtccagatgagcgtgtttagttctgtcagtg 414
Qy 447 tctactgtgttctcgttgcaagtgctgtg---aaactataagcagcagcagcagc 503
Db 415 tctactgtgttctcgttgcaagtgctgtgaaacttataaagcagcagcagcagcagc 474
```

OY	504	gcccgtgaatttgggctctgggggcccgtgcacccaatgcgaagaagatccgaaagcattctgtctctgt	563
Db	475	gctcttgaaatttgggctctgggggcccgtgcacccaatgcgaagaagatccgaaagcattctgtctctgt	534
OY	564	taagtcacaatgtatactgcgtcttgagccattccctggaaactctgaagatataccaagaagtgctgc	623
Db	555	taagtcacaatgtatactgcgtcttgagccattccctggaaactctgaagatataccaagaagtgctgc	594
OY	624	taacattgacatgtatatctcaacccaatgctgcacggcgctggaagaagcctctctacaaccaagac	683
Db	595	taacattgacatgtatatctcaacccaatgctgcacggcgctggaagaagcctctctacaaccaagac	654
OY	684	cgagtcacgaactgtgtctctcttcataaagatacggagagatctctcccaagaactctggagacct	743
Db	655	cgagtcacgaactgtgtctctctcttcataaagatacggagagatctctcccaagaactctggagacct	714
OY	744	cggagataccgggggctctggccaagaagaatatactgctgttaactaccgcgtcccgagacggg	803
Db	715	cggagataccgggggctctggccaagaagaatatactgctgttaactaccgcgtcccgagacggg	774
OY	804	attgatacgaagcttcctctctggggccattttcaagccggtatgtccaaagtatacggagatg	863
Db	775	attgatacgaagcttcctctctggggccattttcaagccggtatgtccaaagtatacggagatg	834
OY	864	ttccaagccatagtctggggtgctcttaacagtgtggtctccagacctccctacatctgggatacgtta	923
Db	835	ttccaagccatagtctggggtgctcttaacagtgtggtctccagacctccctacatctgggatacgtta	894
OY	924	ggttgctctcaactatctatactacaagaagcacgcgaagtgtgtggaatttgtcaagaactt	983
Db	895	ggttgctctcaactatctatactacaagaagcacgcgaagtgtgtggaatttgtcaagaactt	954
OY	984	aacctgcctatactgctgaatcgctcggagagggcggtgttaaccatctgtaaactgtccgggtgc	1044
Db	955	aacctgcctatactgctgaatcgctcggagagggcggtgttaaccatctgtaaactgtccgggtgc	1011
OY	1044	tggacataatggaagaaagctgttgccctctgatacggagagatccctatacgaagcttccatacaat	1101
Db	1015	tggacataatggaagaaagctgttgccctctgatacggagagatccctatacgaagcttccatacaat	1074
OY	1104	gaactacttgaattctcttttggacaacgaattttaaagctccacaataagctctcccaatatagact	1166
Db	1075	gaactacttgaattctcttttggacaacgaattttaaagctccacaataagctctcccaatatagact	1133
OY	1164	aaccagaagaacgaagaaatgagctacccctggagaaagatcaaaacagaagaaatctgttgaagaacttaga	1223
Db	1135	aaccagaagaagaatgagctacccctggagaaagatcaaaacagaagaaatctgttgaagaacttaga	1199
OY	1224	atgtctgcgcgcacgcgaactctgggtgttccaaacgcagagcgatctcctgtagagacgcacacctgag	1281
Db	1195	atgtctgcgcgcacgcgaactctgggtgttccaaacgcagagcgatctcctgtagagacgcacacctgag	1254
OY	1284	gagagctgcgcgatactgagagacgcgaagaacgaacctctggaacaaacgcacatctgcgtctctcttgac	1344
Db	1255	gagagctgcgcgatactgagagagcgaagaacgaacctctggaacaaacgcacatctgcgtctctcttgac	1311
OY	1344	aaacgaaatctgcctgttggagagaagatctctccgaattctgtaaagagagagagagaggggcccgc	1403
Db	1315	aaacgaaatctgcctgttggagagaagatctctccgaattctgtaaagagagagagagaggggcccgc	1374
OY	1404	aaggaactcttccaaactccaaaagaagccaaagaggttcaaaaacgagagatgtaaaaagagaata	1461
Db	1375	aaggaactcttccaaactccaaaagaagccaaagaggttcaaaaacgagagatgtaaaaagagaata	1433
OY	1464	gaacccagaagaggaagaaaggaatatcaaccgaagagagagaaacaaacgaagagagagagccagaata	1523
Db	1435	gaacccagaagaggaagaaaggaatatcaaccgaagagagagaaacaaacgaagagagagagccagaata	1499
OY	1524	gccaaaagggtctcaagggagggagggccaaagtgtgctctgtaatgtgacctctcaagctctctgtcttc	1581
Db	1495	gccaaaagggtctcaagggagggaggttcaaaagtgtgctctgtaatgtgacctctctcaagctctgtcttc	1551

Query Match	50.8%;	Score 818;	DB 24;	Length 1997;
Best Local Similarity	73.4%;	Pred. No. 3.8e-238;		
Matches 1046;	Conservative 0;	Mismatches 380;	Indels 0;	Gaps 0;
QY	126	ggcaccggaggaagtctgttactactacgacgggagtggtgaataatcattatgga	185	
DB	187	ggcggcagaagaagtgctactactatgta19gtgatattgcaattattatattgac	246	
QY	186	caaggccacccaataaagctcacggaatccgcatgagcattatgtgtcaacat	245	
DB	247	cagggtatccatccatggaagctcatatgataccgga1gagcattcaactgtcgtcaatat	306	
QY	246	gtctctccacgaanaatggaaatctatcgcctccacaagccaatgct1gagagatgacc	305	

Db 307 ggtttataccgaaaatgaaatatatagagccttacaagaccacgtctgaagaatgact 366
Qy 306 aagttaccagagctgactacatttaattcttgcgtccatccgtccaaataacatgtcg 365
Db 367 aaatccacagagatgagatatacaagttcttccagataaagacaataataatgltc 426
Qy 366 ggtacacagacagatgacagatattcaacgtttgtgagagactgtccaatattcgaatgc 425
Db 427 ggtacacagacagatgacagatattcaacgttcgagagaagatgtccgtgtttgagaa 486
Qy 426 ctgtttgagttctcagttgtcttactgtgttctgttgcgaagtgtgtgaacttaatt 485
Db 487 ctcttgagttgttcaggtctccacaggtgtgttcagttgtgtgcgtgtgaattaac 546
Qy 486 aagcagcagacagacatgcgtgaattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 545
Db 547 cggcaacaacatgatactgtcaattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 606
Qy 546 gagcaatcgtgtctgttactgtcaatgtatcgtcttgcacatccctgtgaactgtaaag 605
Db 607 gaagcaccaggt 666
Qy 606 tatcaacagaggt 665
Db 667 tatcaacagaggt 726
Qy 666 gctttcaacacagcagcgt 725
Db 727 gctttcaacacagcagcgt 786
Qy 726 ccaggaactgt 785
Db 787 cctgtgacagagagactgt 846
Qy 786 taccgcgtccgagagagactgt 845
Db 847 tttcccaatgagagatgataatgataatgataatgataatgataatgataatgataatg 906
Qy 846 tccaaatgagagatgataatgataatgataatgataatgataatgataatgataatgata 905
Db 907 tcaaaatgagagatgataatgataatgataatgataatgataatgataatgataatgata 966
Qy 906 ctactctgt 965
Db 967 ctgtctgt 1026
Qy 966 gaattgt 1025
Db 1027 gaagt 1086
Qy 1026 cgttaacgt 1085
Db 1087 csgaatgt 1146
Qy 1086 aatgt 1145
Db 1147 aatgt 1206
Qy 1146 aatgt 1205
Db 1207 agtcttcaataatgataatgataatgataatgataatgataatgataatgataatgata 1266
Qy 1206 agtcttcaataatgataatgataatgataatgataatgataatgataatgataatgata 1265
Db 1267 ttattgt 1326
Qy 1266 gaggaagcaccatccgt 1325
Db 1327 gaggaagcaccatccgt 1386
Qy 1326 tgcattgt 1385

Db 1387 tccattcgcacatccagacaacagatagctgtgcataagagtttccagattctgaggt 1446
Qy 1386 gagggagagggggggcgcgaagactcttccaaattccaaaagccaaagaggtcaaaaca 1445
Db 1447 gaaggtgaggggggtcgtgaggaattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1506
Qy 1446 gagatgaaaagagaaagacccagagagagagaaaggaatccacgaagagagaaagacc 1505
Db 1507 attgagagagacaaagaa 1566
Qy 1506 aagggagagagccag 1551
Db 1567 tccaaagacaatagtggtgagaaaacagaccacaaagagagagagag 1612

RESULT 7
AAZ32066
ID AAZ32066 standard; DNA; 1985 BP.
XX
XX AAZ32066;
AC
XX
XX 10-JAN-2000 (first entry)
DT
XX
XX Human METH2 related EST U31814.
DE
XX
XX Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;
KW cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
KW angiogenesis inhibitor; abnormal wound healing; inflammation;
KW rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
KW diabetic retinopathy; macula degeneration; haemangioma; detection;
KW arterial-venous malformation; immune deficiency; ss.
OS
XX
XX Homo sapiens.
PN
XX W09937660-A1.
PD
XX 29-JUL-1999.
XX
XX 22-JAN-1999; 99WO-US01313.
PF
XX
XX 23-JAN-1998; 98US-0072298.
PR
XX 28-AUG-1998; 98US-0098539.
XX
XX (IRUE/) IRUELA-ARISPE L.
PA (HAST/) HASTINGS G A.
PA (RUBE/) RUBEN S M.
XX
XX IrueLA-Arispe L, Hastings GA, Ruben SM;
PI
XX WPI: 1999-590684/50.
DR
XX
XX New isolated metalloprotease thrombospondin polypeptides, useful for
PT treating hyperproliferative disorders, cancers or autoimmune disorders
PT
XX
XX Disclosure: Page 444-446; 457pp; English.
PS
XX AAZ32000 and AAZ32001 encode, and AAY49501 and AAY49502 represent, human
CC metalloprotease thrombospondin (METH) proteins METH1 and METH2
CC respectively. METH1 and METH2 have been found to be potent inhibitors of
CC angiogenesis both in vitro and in vivo. They can be used for treating
CC cancer and other disorders related to angiogenesis including abnormal
CC wound healing, inflammation, rheumatoid arthritis, psoriasis,
CC endometrial bleeding disorders, diabetic retinopathy, some forms of
CC macula degeneration, hemangiomas, and arterial-venous malformations.
CC They may be useful in treating deficiencies or disorders of the immune
CC system, by activating or inhibiting the proliferation, differentiation,
CC or mobilisation (chemotaxis) of immune cells. The etiology of these
CC immune deficiencies or disorders may be genetic, somatic, such as
CC cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
CC toxins), or infectious. They can also be used to treat inflammatory
CC conditions, both chronic and acute conditions. The products can also be
CC used for detection and diagnosis. AAZ32002 to AAZ32080, and AAY49503 to

CC AMV9311 represent sequences given in the exemplification of the present
CC invention.
XX
S0 Sequence 1985 BP; 626 A; 360 C; 454 G; 545 T; 0 other;

Query Match	50.0%	Score 805.2	DB 20	Length 1985
Best Local Similarity	72.8%	Pred. No. 3e-24		
Matches 1038	Conservative	0	Mismatches 388	Indels 0
				Gaps
QY	126	gaccacccggagaaagtcgttactactaagacgggagctgttgaaattactatctatgga	185	
Db	223	ggcgcaaaaaaaagcgtctactactactaagcggatgttggaaattattatctatgga	282	
QY	186	caaaagccaccaatgatgaagcctcacccgaatccgcatgactccaataattgtctccaact	245	
Db	283	caaggtcatcccaagaagcctctatgaatccgcatgaccataactgtctgttaaatcat	342	
QY	246	ggtctccacgaaaatgagaaatctatcgccccaagaacgaatgctgaagatgtgacc	305	
Db	343	ggtctatacagaanaaatgagaaatataatggccccaataagccaacgtccgaagaatgaca	402	
QY	306	aagtaacacagcgatgactcatcaatattcttgcgtccatccggtccagatacatgtcgc	365	
Db	403	aaatatacagtgatgtggtatatacaaatcttactaggtctacataagaacgaatacatgtct	462	
QY	366	gagtaacgcaagcagatgacagagatccaagtttggggaagctgtccagattcatgctgc	425	
Db	463	gagtaataagcagatgacatataattatgttgggaagaatgttccgacgtttgatgga	522	
QY	426	ctgtttgagttctgtcagttgtctactgtgtgttctctgtggaagtgtctgtgaactaat	485	
Db	523	ctcttgaagttcttgcagctctcaactggggttcaggtctgtcggagctgtgaagttaac	582	
QY	486	aagcagacagacggaacacgcgcgcgaattggcgctggggggcctgcacatccaagaagtc	545	
Db	583	cgacaacagactgatatgtgttattgttggcgggagatacatcatgtctaagaatcac	642	
QY	546	gaagacatcgtgtctctgttaccgaatgatataatgcgttcggacatccctggaactgtaaag	605	
Db	643	gaagcatcaagatctctgttaccgaatgatataatgtgtctgcacatcccttgaattactaaag	702	
QY	606	tatcaccaagggggtgcgttaccatgtgacatgtgattcacacacatggtgaagcgttgaaag	665	
Db	703	tatcatcagagagcttatatatattgtatataagataattcacatcagtgatgtgttgaaga	762	
QY	666	gcctctcaacacacgacggcggtcatgactgtgtccttcaataagtaagagagtaacttc	725	
Db	763	gcctttatatacaacagatctgttgaatgacggtatcatccataaataatataatggaatacttc	822	
QY	726	ccaggaacatgggacacacggaatcacggggtgcgggaagaagcaagatattatcttcaac	785	
Db	823	ccgtgcacagagagactgtgaaggtatgtgtgcgtcggaagaagcaataactatgtctgcaat	882	
QY	786	taccgcgtccgagagacggtatgtatgacagagctcatagtagccaatttccaacgggtcatg	845	
Db	883	tttccaatgtgtgtatgtatagatagatgtatggtctcatatggaacgaattttaagcccttattc	942	
QY	846	tccaagaataatgagatgttccacgacctagtgcggtgtcttcaagatgtgtcccaaccc	905	
Db	943	tcaaaagtgtatgagatgatatcaactatgtctgtgtattacagtgtgtgtgcagaactca	1002	
QY	906	ctatctgggagatcggtatgagttgcttccaatctatcatcaagaagacacgcaagtggtg	965	
Db	1003	ttatctgttatagacatgggtgttttccaactctacacagtaagaagttcatgtctaagtgta	1062	
QY	966	gaatttctcaagagccttcaacgcctcatgcatgcatgtctggggcggtgtgttaacccct	1022	
Db	1063	gaagttgtaaaaaacttctaactcaactacatcagatgtcttggggaggtgtgtctacaacatc	1122	
QY	1026	cgtaagctgcccgggtgtcgtgacatatgagacagctgtgtgcccctgtgatacagagatccct	1085	
Db	1123	cgtaagtgtctcgatgttggacatctggaacatggaacgcaggtgtgacctgtatgttaagctcc	1182	

QY	1086	aatgaagctccatcatgaatgaactatttgatattcttggaccagattccaagctcaatc	1143
QY	1183	aatgattgtgccatatataatgattactttagtatttctggaccagattccaatgtcattt	1242
QY	1146	agctctccaatatgactaaacccaagaacagaaatgagttacctggaagaagatcaaacagca	1205
Db	1243	agctctcaacaatagacaacaccgaacacccccaagatatatgtaaaagataaacacgct	1302
QY	1206	ctgttttggaaccttaagaatgtctgcgcacgcacctggggtccaaagcagatccct	1265
Db	1303	tgttttggaataatttgcgatgttlaacctcatgcacccctggtgtccagttgtaagctattcca	1362
QY	1266	gaggagcccatccctgtgagagatgtgcgatatgagagcaagaacgacccctgaacgcgcatc	1325
Db	1363	gaagatgtctgtcttcatagaacaatgtagagatgtgaagatccagaacaagagaatt	1422
QY	1326	tcgatctctctctctgtgcaaacagatgtgcctgtggaagaagatctccgattctgaagag	1385
Db	1423	tctattcgagcatcagaacaacggtatcgtcctgtgtagaagaattctcattctcgtagat	1482
QY	1386	gaggagagagggggcgcgcgaagaactcttccaaattcaaaagccaaaggagttcaaca	1445
Db	1483	gaaggagagaaggaggtcgaagaagaatgtgtcgtcatcataagaagaggagcaagaagctaga	1542
QY	1446	gaggatgtaaaaagagaaagaccccgagagagagaaagaaagaaatccagaaagagaaacc	1505
Db	1543	attgaagaagaataagaagaacaacagagacaaaaaaaacacgacttaaggaaagataaa	1602
QY	1506	aaggagagagagccagaagccaaagggtcagaaggagagccaggt	1551
Db	1603	tccaaggacacagtggtgtgaaaaaacgataccaagaagaaaccaat	1648
RESULT 8			
AAC90323			
ID	AAC90323 standard; DNA; 1985 BP.		
XX	AAC90323;		
AC	19-MAR-2001 (first entry)		
XX	U31814 cDNA clone.		
DE	MEH1; metalloprotease; thrombospondin; angiogenesis inhibition;		
KM	cancer therapy; benign tumour; ocular angiogenic disease;		
KM	rheumatoid arthritis; psoriasis; wound healing; endometriosi;		
KM	vasculodermis; granululation; hypertrophic scar; nonunion fracture;		
KM	scleroderma; trichoma; vascular adhesion; myocardial angiogenesis;		
KM	coronary collateral; cerebral collateral; arteriovenous malformation;		
KM	ischemic limb angiogenesis; Osler-Weber syndrome; wound granulation;		
KM	plaque neovascularisation; telangiectasia; haemophilic joint; EST;		
KM	angiolipoma; fibromuscular dysplasia; expressed sequence tag;		
KM	Crohn's disease; atherosclerosis; birth control; ss.		
XX	Unidentified.		
OS	WO200071577-A1.		
XX	30-NOV-2000.		
PD	25-MAY-2000; 2000WO-US14462.		
XX	25-MAY-1999; 99US-0318208.		
XX	20-JUL-1999; 99US-0144882.		
PR	10-AUG-1999; 99US-0147823.		
PR	13-AUG-1999; 99US-0373658.		
PR	22-DEC-1999; 99US-0171503.		
PR	22-FEB-2000; 2000US-0183792.		
XX	(HUMA-) HUMAN GENOME SCI INC.		
PA	(SMIK-) SMITHKLINE BEECHAM CORP.		
PA	(BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.		

PA (IRUE/) IRUELA-ARISPE L
PA (HAST/) HASTINGS G A.
PA (RUBE/) RUBEN S M.
PA (JONA/) JONAN Z L.
PA (TRUL/) TRULLI S H.
PA (FORN/) FORNWALD J A.
PA (TERR/) TERRETT J A.

PI Irueala-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trullii SH;
PI Fornwald JA, Terrett JA;

PT METH1 and METH2 polynucleotides and encoded polypeptides, used to
PT inhibit angiogenesis in the treatment of disorders such as cancer,
PT rheumatoid arthritis and psoriasis -

PS Claim 14; pages 745-746; 768pp; English.

The present invention relates to human METH1 and METH2, (ME for metalloproteinase and TH for thrombospondin; see BAB50002 and BAB50003). The present sequence is an expressed sequence tag (EST) for METH. METH can be used for inhibiting angiogenesis in an individual, and for treating cancer, benign tumours, an ocular angiogenic disease, rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis, vasculodermatitis, granulatiions, hypertrophic scars, nonunion fractures, scleroderma, trachoma, cerebral adhesions, myocardial angiogenesis, coronary collaterals, cerebral collaterals, arteriovenous malformations, ischaemic limb angiogenesis, Osler-Weber syndrome, plaque neovascularisation, telangiectasia, haemophilic joints, angiofibroma, fibromuscular dysplasia, wound granulation, Crohn's disease or atherosclerosis. METH can also be used in birth control. METH can also be used in diagnostic methods for the prognosis of cancer.

Sequence 1985 BP; 626 A; 360 C; 454 G; 545 T; 0 other,

Query Match	50.0%	Score 805.2	DB 22	Length 1985
Best Local Similarity	72.8%	Pred. No. 3e-234		
Matches 1038; Conservative	0;	Mismatches 388;	Indels 0;	Gaps 0

OY	126	ggccccgggagaaagcctgttactactaagaagggatgtgttgaaattactattaga	155
Dp	223	ggcgcgcaaaaaaaagctctgttactactaagaaggtatattgaaattatttagga	262
OY	186	caagggcaacccaatgaaagccctccacgcatccgcatgtactataattgtgcgcacaat	245
Dp	283	cagggtcatcccatgaagccctcatagatatcgatgcccataactgtgcgttaaatat	342
OY	246	ggtcttaacgaaaaatggaattatctgcctccaagaagccaattgtcgagatgacc	305
Dp	343	ggctatacagaaaaaatggaattatattagggcccatlaagccactgcgcgaagaattaga	402
OY	306	aagtaaccaagcgatgtgactaatattcttggcgccatccgttcagataaacaatgtg	365
Dp	403	aaatataccagtgatgtgatatatactaaattctcaagctcaataagaacagataaacaagtct	462
OY	366	gagtaacgcaagcgatgacagtgagattcaacgcttvgtgagagctgtccagatattcgatg	425
Dp	463	gagtatagtaagcagatgcatattatttaactgttvgagaaagattgtccagcggtttagtga	522
OY	426	ctgtttagttcttcaagtggtgtcactgggtgttctctgtgcaagtggttggaaactaat	485
Dp	523	ctctcttgagttcttgcagcctccaacggtcggtttcaattgtcgggaagcttggaaagttaac	582
OY	486	aagcagcagagacgacatccgcgttgaatttggcgctvgggcgcttgcacacatgcanaagttcc	545
Dp	583	cgacaacagactgattatgctgttgaatttggcgctvggagattacaatcatgtcaagaataac	642
OY	546	gagcgatctgtgctctgtttagcgtcaatagatatgctcttggccatccctggaactgtcaag	605
Dp	643	gaacgacatcaggaattcttcaagttaaagatatgtgttgccttcactccttgattactaaag	702

OY	606	taaaccaagaggggctgcgtacattgacattgatactccaccagggggaagcggttgaaaag	665
Db	703	tatcatcagaagagcttataataatattgataatgataatccatcagatggtgagtgctgtgaa	762
OY	666	gaccttcacacccaagggacccgggtcagtaagctgtgtcccttcataagtatgaaagtac	725
Db	763	gcttttatatacaaaagatctgtgtaatgaaagcgtatactccataaataatggaataactt	822
OY	726	ccaggaactggggacctaagggatbacccgggctgcgaagaagtatattgctgttaac	785
Db	823	ccctgcacaaagagactctgagggatattgtgtcgtgaaaaagcaataactatgtctcat	882
OY	786	taccggctccggaagacgggatttgatgagcaggtctcatgaaagccattttaagccggct	845
Db	883	tttccaatgtgtgtgtgtaataagatgataatgagtggaagatattaaagcttatactc	942
OY	846	tccaaataatagagatgattccacagctagtcggtgtgtcttcacagtggtgcacagctcc	905
Db	943	tcaaagtgatgaggaatgatactcaactagtgctgtgtatctacagtggtggtgcagctca	1002
OY	906	ctatctgggagctcggtatgagttgtcttccaatctatactacaagagacacggcaagtgtg	965
Db	1003	ttactcgtgtagatagacgtgggtgtgttccaactcaacagtaaaagtcataatgtgtga	1062
OY	966	gaatttgccaagagctttaacctgaacctgactatgctgtgagggcggtgtgtacaacctt	1025
Db	1063	gaagtgttaaaaaacttttaactctacatactatctgtctgtgagggaggtgtgtctacaactc	1122
OY	1026	cgtaagcttgcggcggtgctgacatactgaaacagcgtgtgtgcccctgatacggaaagctcc	1085
Db	1123	cgtaatgtgtctcgatgtgttgacatatgaaagacgtgcagctgtgcccttgattgtagatctcc	1182
OY	1086	aatgagcttccatacaaatgatactactttgataacttggacagacagatttcaagctccacat	1145
Db	1183	aatgagttgcataataatgatagttacttggatttggacagacttcaaaactgacatat	1242
OY	1146	agtccttccataatgactataaccgcaaacacgaataggttaccttggaagaagatcaaaagga	1205
Db	1243	agtccttccaacaactgaaacaacccagaacacactccagaatatatgaaagaagataaaacagct	1302
OY	1206	ctgtcttgagaaccttgaatagctgtgcgcgcagcacccgtgggttccaacgaagcgtactct	1265
Db	1303	ctgtttgaaaaatttgcgcatgtaactctcatgacccctggtgtccagatgacgaactatcca	1362
OY	1266	gaggaagccalccccttgagggaggttggtgcagatgaaagacagaaacccctgacaaagcgtcat	1325
Db	1363	gaagatgctgttcaatgaaagccagttgagatggaagtggagaagatccagacaagaagaatt	1422
OY	1326	tcgatactgcctctctgacaaacgaatgctgctgtgaggaagattctccgattctgaaagag	1385
Db	1423	tcctatctcgatcatcagaacaaagcggataagcttctgtatgaaagaattctccagattctgaggt	1482
OY	1386	gagggagaggggggcccgaagaactcttccaacttcaaaaaaagccaaagaggttcaaaa	1445
Db	1483	gaagaggaagaaagaggttgcgaagaatgtgtgcgtgacatataaagaagggcagaagaagctaga	1542
OY	1446	gagagatcaaaaagaaagaccccaagcggagaaagaaagagatcacogaagagagagaanaac	1505
Db	1543	attgaagaagaataagaaagaacacagcggacgaanaaaaaacagcgttaaggaagaagataa	1602
OY	1506	aaggaagagaagcccaagaagccaaagggtgtcaaggaagagcccaagt	1551
Db	1603	tccaagagacacagtggtgtgaaaaaacaagataccaaagggagacaaat	1648

RESULT	9
ACC89555	
ID	ACC89555 standard; DNA; 1985 BP
XX	
AC	ACC89555;
XX	
DT	08-MAR-2001 (first entry)
XX	

DE Human histone deacetylase HDAC-2 coding sequence.
XX
KM Histone deacetylase; HDAC-1; HDAC-2; HDAC-3; HDAC-4; HDAC-5; HDAC-C;
KM HDAC-D; cell cycle; tumorigenesis; cancer; inhibitor; antisense;
KM gene therapy; ds.
XX
OS Homo sapiens.
PN MO200071703-A2.
PD 30-NOV-2000.
PF 03-MAY-2000; 2000MO-IB01252.
PR 03-MAY-1999; 99US-0132287.
PA (METH-) METHYLENE INC.
PI Macleod AR, Li Z, Besterman JM;
XX WPI: 2001-016407/02.
XX P-PSDB; AAB49955.
PT Antisense oligonucleotide that inhibits expression of a histone
PT deacetylase, useful for treating and/or alleviating the symptoms of
PI neoplasia, or for inhibiting neoplastic cell growth in an animal -
XX
XX
PS Disclosure; Page 53; 125pp; English.
CC The present invention provides inhibitors of histone deacetylase enzymes
CC such as HDAC-1, HDAC-2, HDAC-3, HDAC-4, HDAC-5, HDAC-C and HDAC-D. These
CC inhibitors may be antisense strands or they may be compounds identified
CC by contacting the enzyme with the compound and measuring the resulting
CC enzyme activity. These inhibitors are useful for treating cancers and for
CC identifying which histone deacetylase is involved in a neoplasia.
XX
XX Sequence 1985 BP; 626 A; 360 C; 454 G; 545 T; 0 other:
SO

Query Match 50.0%; Score 805.2; DB 22; Length 1985;
Best Local Similarity 72.8%; Pred. No. 3e-234;
Matches 1038; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

QY 126 ggcaccggaggaagtcgtctactactagacgggagtggtggaattactattatga 185
DB 223 ggcggcgaataaaagtcgtctactactagacgggagtggtggaattactattatga 282
QY 186 caaggccaccatgaagctcaccgaatccgcatgactcataattgtgttcaactat 245
DB 283 cagggtcatcccatgaagctcaccgaatccgcatgactcataattgtgttcaactat 342
QY 246 ggtctctaccgaataatggaatctatcgccctcaaaagccatgctgaagagatgac 305
DB 343 ggcctatacagaataatggaatctatcgccctcaaaagccatgctgaagagatgac 402
QY 306 aagtcacagcgatgactactaattcttgcgtccatccgtccacagatacatgtcg 365
DB 403 aatctcagcgatgactactaattcttgcgtccatccgtccacagatacatgtcg 462
QY 366 gagtcacagcgatgactactaattcttgcgtccatccgtccacagatacatgtcg 425
DB 463 gagtcacagcgatgactactaattcttgcgtccatccgtccacagatacatgtcg 522
QY 426 ctgtttggttcgtcagttgtctactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 485
DB 523 ctgtttggttcgtcagttgtctactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 582
QY 486 aagcgcagcagcgatcgcggtgaattggtgtgtgtgtgtgtgtgtgtgtgtgtgt 545
DB 583 cgacacacgagcgatcgcggtgaattggtgtgtgtgtgtgtgtgtgtgtgtgtgt 642
QY 546 gaggcacggt 605

DB 643 gaagcatcaggatctgttaagcttaagatattgtgtgtgtgtgtgtgtgtgtgtgt 702
QY 606 tatccacgaagggt 665
DB 703 tatccacgaagggt 762
QY 666 gctctctacacacgagccgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 725
DB 763 gcttttatacacaagatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 822
QY 726 ccagggaactggggagcctaaaggaatccgggtgtgtgtgtgtgtgtgtgtgtgtgt 785
DB 823 cctggcagcaggaactggggagcctaaaggaatccgggtgtgtgtgtgtgtgtgtgt 882
QY 786 tacccgctccgagcagggatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 845
DB 883 ttccaagt 942
QY 846 tccaaagtaattgt 905
DB 943 tcaaaagt 1002
QY 906 ctatctggggatcggttaagttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 965
DB 1003 ttatctgt 1062
QY 966 gaattgtcagaagcttcaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1025
DB 1063 gaattgtcagaagcttcaactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1122
QY 1026 cgttaagttgccgggt 1085
DB 1123 cgttaagttgccgggt 1182
QY 1086 aatgagcttcacataagcacttgaacttgaacttgaacttgaacttgaacttgaact 1145
DB 1183 aatgagcttcacataagcacttgaacttgaacttgaacttgaacttgaacttgaact 1242
QY 1146 agtccttccaatatactactacacagacagaaatgagtaactgtgagaagatcacaagcga 1205
DB 1243 agtccttccaatatactactacacagacagaaatgagtaactgtgagaagatcacaagcgt 1302
QY 1206 cgttttgagaacctagaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1265
DB 1303 tctgttgaataattgt 1362
QY 1266 gaggaagccatccctggaagaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1325
DB 1363 gaagatgtcgttcatgaaagacagtggaatgaaatgaaatgaaatgaaatgaaatgaaat 1422
QY 1326 tcatctgtctctctgacacaaagaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1385
DB 1423 tcatctgtctctctgacacaaagaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1482
QY 1386 gaggaagaggggt 1445
DB 1483 gaggaagaggggt 1542
QY 1446 gaggaatgaaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaag 1505
DB 1543 attgaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaa 1602
QY 1506 aaggaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaa 1551
DB 1603 tccaaggaacacgt 1648

RESULT 10
ABL01897
ID ABL01897 standard; cDNA: 2145 BP.
XX
XX ABL01897;
XX

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 170.
 DE Drosophila: developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX Drosophila melanogaster.
 OS
 PN MO20171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PE 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 BK WPI; 2001-656860/75.
 XX P-PSDB; ABB57793.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Claim 1; SEQ ID NO 170; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX Sequence 4755 BP; 1321 A; 1161 C; 1023 G; 1250 T; 0 other;
 SO

Query Match 39.8%; Score 640.6; DB 23; Length 4755;
 Best Local Similarity 72.5%; Pred. No. 6.6e-184;
 Matches 829; Conservative 0; Mismatches 314; Indels 0; Gaps 0;

QY 266 aatcatatgacccatcaagaacgaatgctgagagatgacccaagtaccacagcatgacta 325
 DB 2668 ATTCTAGCGCTCCCAATAAGCCACTGCCGATGAGTACCAAGTCCACTCGGAGAGACTA 2809

QY 326 catlaaatctctgcctcaccatcgcgcagatacaacatgctcgcagatcgcagacagatgca 385
 DB 2808 CGTCCGGTCTTTCGATCATTCGCGCGGACACATGTCGAGTCAACAAAGCATATGCA 2749

QY 386 gagattcaacgttggtagaactgtccagatattcgaatggcgtgttgatgttcagtt 445
 DB 2748 GCGTTTCAATGTGCGGAAGATTGTCGCTTTCGATGAGATATGAGATTTCGCAACT 2689

QY 446 gtctactgtgtgtctgtggaagtgtgtgtaacttaataagcagaagagatcgc 505
 DB 2668 CTCCGCCGAGAGATCCGATGCGCGCGTAAACGTGAATGAAGCCTCGGAGATCTG 2629

QY 506 cgtgaattggcgtgggagcctgcaccatcgaagaagtcgagatctgcttctgtta 565
 DB 2628 CATCATATGGGGCGGTGGGCTGATCAGCCCAAAAATCGAAGCTCTGGGCTTGTGCTA 2569

QY 566 cgtcaatgatatcgttctgtgcatctcgtgaactgctaaagatcacccagaggtgtc 625
 DB 2568 CGTCAACGACATTTGTTTGGGAATTCGTGAATCTGTAATATACCATCATGCGGTCTCTA 2509

QY 626 catgacattatataatccatcgtgtgacggtgtggaagagcctctacaccagagacg 685
 DB 2508 CATAGATATAGACGTCCATCCACGGCGATGGCGTGGAGAGCGTCTATACCCAGATCG 2449

QY 686 gttcaatgactgtcccttcaatgaatgatgagatctcccaagaactgggagactacg 745
 DB 2448 TGTGATGACTGTACAGTTCACAAAGTATGAGATATTTCCGGGCACTGGCGATCTCG 2389

QY 746 ggalacggggctggaagaagaatattatgctgttaactaccgcctccgagacgggat 805
 DB 2388 AGACATTTGGCCCGCGCAAGGAAAGTACTATGCGGTGAATATACCCCGCGATGGCAT 2329

QY 806 tgatgacgaatcctatgtaggccaatttcaagccggtatcctccaaagtatgagatgt 865
 DB 2328 GGATGATATATGCGTACGAGAGCATTTTGTGCCCATTTATTCAGCAAGATGATGAACAT 2269

QY 866 ccagcctagtcggtgtgtcttaccagtgtgtcgaagccctatctgggagatcgttag 925
 DB 2268 CCAAGCCGCGACCGTGTGTGTGCACTGTGGCCGCAATTCCGACTGGCGATCGGTTAGG 2209

QY 926 ttgcttcaatcatalcatalcaaaagacgcaagtggtgtgaattgttcaagagctttaa 985
 DB 2208 CTGCTTCAATCTCACCGTCAAGGGTCAAGGCAAGTGGTGAGTTCGTGAAGAAATATAA 2149

QY 986 cctgtctatgtctgtagtgcgggagagcggtgtgttaccacatcgttaacgttgcgggtgt 1045
 DB 2148 CTTGCCATTCTGATGATGTCGGGTGTGTTATTCATTCGTATGTTATCCCGCTGCTG 2089

QY 1046 gacatatgacagctgtgcccctgtgatacgagataccctaatgagcttcatacaatga 1105
 DB 2088 GACCATATGAGACCTCCGTTGCACTGCGGTGAGATACCAACGAACTGCCCTACAAAGCA 2029

QY 1106 ctacttgaatacttggacaagaatttcaagctccacatcagtcctccaatatgactaa 1165
 DB 2028 TTACTTCAAGTACTTTGGGCGCGATTTTAAGCTGCACATTTAGTCCACCAATATGACGAA 1969

QY 1166 ccagacgaacgaatgtagtctccctgggagaagaatcaaaagagactgtttggaaccttaagat 1225
 DB 1968 TCAGATATACATCCGATGCTCGGAGAGATCAAGAACCTGTCTGTCGAGACTGCGCAT 1909

QY 1226 gctgcgacgacacccctgggtgtccaaacgacgagatctccttgaagacacccctgaaga 1285
 DB 1908 GCTGCTCACCGCTCCGGGCGTTCMAATCCAAAGCATTTCCGAGATGCCATTCMAAGATGA 1849

QY 1286 gattgagatgagagacgaagaagaccctgacaagcgcatctgatcgtcctctgacaa 1345
 DB 1848 GTCCGACGACGACGACAAAGTTCGACAAAGATGATGCTGCGCGACAGACGACAAAGACAA 1789

QY 1346 agcaattgcctgtgaggaaggtctccgattctgaagaagaagagagggggcgga 1405
 DB 1788 GCGCATTTGTGCCGGAAGACGATCTCCGATTCCGAGATGAGGCGAAGCGGTGCGAG 1729

QY 1406 gaa 1408
 DB 1728 GGA 1726

RESULT 12
 ID AAX90840
 ID AAX90840 standard; DNA: 1943 BP.
 AAX90840:
 13-JAN-2000 (first entry)
 Maize histone deacetylase-4 DNA.
 Maize histone deacetylase: family 1, ZmHD1; promoter regulator; promoter;
 RNA polymerase II; transcription; plant transformation; heterochromatin;
 disease resistance; chromatin assembly; gene activity; toxin screening;
 pathogenicity; disease response promoter; ds.
 Zea mays.

XX	Key	Location/Qualifiers
FT	CDS	51..1610
FT		/tag=a
FT		/product="Maize histone deacetylase"
XX		W09951731-A2.
XX		14-OCT-1999.
XX		02-APR-1999; 99WO-US07370.
XX		03-APR-1998; 98US-0080563.
XX		(PION-) PIONEER HI-BRED INT INC.
XX		Baldwin DA, Briggs SP, Crane VC;
XX		WPI; 1999-611038/52.
XX		P-PSDB; AAY28800.
PT		New deacetylase genes, used for producing transgenic plants which have
PT		increased disease resistance
PS		Claim 1; Page 60-63; 87pp; English.
XX		The present sequence encodes a maize histone deacetylase. This DNA
CC		belongs to family 1, zmhD1 and appears to be a regulator of promoters for
CC		RNA polymerase II, for transcription of genes. The nucleotide sequence
CC		can be used to transform plants and increase disease resistance by
CC		altering the levels of heterochromatin, altering chromatin assembly, and
CC		gene activity of the transformed plants. Additionally, compositions find
CC		use in screening for toxins that affect pathogenicity and in determining
CC		which disease response promoters are regulated by histone deacetylase.
XX		Sequence 1943 BP; 510 A; 498 C; 497 G; 438 T; 0 other;
SQ		
	Query Match	31.2%; Score 502.6; DB 20; Length 1943;
	Best Local Similarity	65.5%; Pred. No. 4.1e-142;
	Matches 768; Conservative	0; Mismatches 399; Indels 6; Gaps 2;
OY	134	gaggaaagctctgtactactacgacgaggagtggtggaatactattatgagcaagcca 193
DB	122	gggggagctgtgtactactacgacgaggagtggtggaatactattatgagcaagcca 181
OY	194	cccaatgaagctcaccgacatcgcatgacataattgtctgctcaataatgagctcta 253
DB	182	tccgtagaagccgacgacatcgcatgacataattgtctgctcaataatgagctcta 241
OY	254	ccgaaatgaatctatcgctctcaacaaagccaatgctgagagatgacccaatgaca 313
DB	242	caaccagatgagtggtacacgcccacccgcccgcgcgcgcgcgcgcgcgcgcgc 301
OY	314	cagcgatgactaatattcttgctccatccgctccagtaataatgctggagtaag 373
DB	302	cgcgcagcactacatacattctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 361
OY	374	caagagatgacagatcaacgcttggtgagagctgacagatctcgaagctgtttga 433
DB	362	ccgctgtcgaagcgttcaacgcttggtgagagctgacagatctcgaagctgtttga 421
OY	434	gtctgtcagatgtgtctactgtgtgtctgtggaagtgctgtggaactaataagcaga 493
DB	422	ctctgtcagacattgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 481
OY	494	gacgagacatcgccggaatgtgctgctgagcctgacacatgcaagaagatccgagcgc 553
DB	482	---tgacacgacatcaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 538
OY	554	tggctctgttcaatgatatgctctgtgacatccggagactgcaataatgatacaca 613
DB	538	gggtctgtctgctgacatgacatgctgtctgacatccgagctgtgtcaatgatacaga 598

OY	614	gagggtgtctatcatgacattgatactacacatgtgtgacgctgtggaagagcctcta 673
DB	599	gagagttctgtatgtcatatcatatccacacatgtgtgagtgagtgagtgagtgagtg 658
OY	674	caccacgagccggtgacatgactgtgtctcttcaataatgataatgagatcttccaggaac 733
DB	659	cacacagataggtgtatgactgtgtctcttccacacatgtgtgtgtgtgtgtgtgtgtgt 718
OY	734	tgggactacgaggtatccaggggctgtgcaagaagatataatgtgtgtgtgtgtgtgtgt 793
DB	719	agggatataccgtgacatgt 778
OY	794	ccgagaggtatgt 853
DB	779	agatgt 838
OY	854	aatgagatgttccagctgt 913
DB	839	tatgaggttccgctgt 898
OY	914	ggatgt 973
DB	899	ggatgt 958
OY	974	caagagcttaaccctgt 1033
DB	959	gaggtcttcaacgttccatgt 1018
OY	1034	tgcgcgt 1093
DB	1019	tgcagcgt 1078
OY	1094	tccatcaatgt 1153
DB	1079	gctgt 1138
OY	1154	caatgt 1213
DB	1139	taacatgt 1198
OY	1214	gaactgt 1270
DB	1199	taactgt 1258
OY	1271	cgcacccctgt 1303
DB	1259	agaatgt 1291
RESULT 13		
AA90839		
ID	AA90839 standard; DNA; 2019 BP.	
XX		
XX	AA90839;	
DT	13-JAN-2000 (first entry)	
XX		
DE	Maize histone deacetylase-3 DNA.	
XX		
KW	Maize histone deacetylase; family 1, zmhD1; promoter regulator; promoter;	
KW	RNA polymerase II; transcription; plant transformation; heterochromatin;	
KW	disease resistance; chromatin assembly; gene activity; toxin screening;	
KW	pathogenicity; disease response promoter; ds.	
OS	Zea mays.	
XX		
XX	Key	
FT	Location/Qualifiers	
FT	CDS	
FT	140..1459	
FT	/tag=a	
FT	/product="Maize histone deacetylase"	
XX		
PN	W09951731-A2.	

XX 14-OCT-1999.
PD
XX 02-APR-1999; 99MO-US07370.
PF
XX 03-APR-1998; 98US-0080563.
PR
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Baldwin DA, Briggs SP, Crane VC;
PI
XX WPI: 1999-611038/52.
DR P-PSDB; AAY28799.
XX
XX New deacetylase genes, used for producing transgenic plants which have
PT increased disease resistance
PS
XX Claim 1; Page 55-58; 87pp; English.
XX
XX The present sequence encodes a maize histone deacetylase. This DNA
XX belongs to family 1, ZmHDI and appears to be regulator of promoters for
XX RNA polymerase II, for transcription of genes. The nucleotide sequence
XX can be used to transform plants and increase disease resistance by
XX altering the levels of heterochromatin, altering chromatin assembly, and
XX gene activity of the transformed plants. Additionally, compositions find
XX use in screening for toxins that affect pathogenicity and in determining
XX which disease response promoters are regulated by histone deacetylase.
XX
XX Sequence 2019 BP; 510 A; 532 C; 528 G; 449 T; 0 other:
XX

Query Match 31.1%; Score 501.8; DB 20; Length 2019;
Best Local Similarity 65.3%; Pred. No. 7.3e-142;
Matches 753; Conservative 0; Mismatches 397; Indels 3; Gaps 1;

QY 121 cgcagggaccccggaagatctgttactactacgacgggagtttgaaattactatt 180
DB 189 cggaggggtcgaagcgcgctgtctacttctacgacggaggttggaactactact 248
QY 181 atggaacggccacccaatgaagctcaacgacgacgacgacgacgacgacgacgac 240
DB 249 acgggcaaggccacccgagtgaaagccgacgacgacgacgacgacgacgacgacgac 308
QY 241 actatggtctctaccgaaanaatgaaatctatcgccctccacaaagccaatgctgagaga 300
DB 309 gctacggtcctctccgacagatgagtgcttcgcccctccacccctgcccgcgcgacgc 368
QY 301 tgaccaagtaccacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 360
DB 369 tctgcgctctccacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 428
QY 361 tctgcgaggtaccgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 420
DB 429 agcagggaccacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 488
QY 421 atgacctgttgcagtgctgcagtgctgcagtgctgcagtgctgcagtgctgcagtgctgcag 480
DB 489 acggtctctacagtgctgcagtgctgcagtgctgcagtgctgcagtgctgcagtgctgcag 548
QY 481 ttaataaagcagacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 540
DB 549 tcaacacatggcaca--tgatatcgcatcaatggcgcgagctccacacgacgacgacgacgac 605
QY 541 agtcgagagcatctgtctgttgcagacgacgacgacgacgacgacgacgacgacgacgacgac 600
DB 606 agtgcgagagcctccggtgttgcagacgacgacgacgacgacgacgacgacgacgacgacgac 665
QY 601 taaagtlataccacgaggtgctgtacatgtacatgtacatgtacatgtacatgtacatgtacat 660
DB 666 tcaagtlataccacgaggtgctgtacatgtacatgtacatgtacatgtacatgtacatgtacat 725
QY 661 aagaagccttcaacacgaggtgctgtacatgtacatgtacatgtacatgtacatgtacatgtacat 720
DB 661 aagaagccttcaacacgaggtgctgtacatgtacatgtacatgtacatgtacatgtacatgtacat 720

DB 726 agggagccttttatccacacgacgaggtgacgacgtccatccacaaagtttgagatt 785
QY 721 acttcccaagaaactgggagacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 780
DB 786 attccctgggacagggagacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 845
QY 781 ttaactaccgctccgagacgaggtatgtacgacgacgacgacgacgacgacgacgacgacgac 840
DB 846 tgaatgttccctggagcagtgatgtacgacgacgacgacgacgacgacgacgacgacgacgac 905
QY 841 tcatgtccaaagtaatggagatgttccacgacgacgacgacgacgacgacgacgacgacgacgac 900
DB 906 taatggcacaaggtgatgtgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 965
QY 901 actccctactgtggagatcggttaggtttgcttcaatctatctatccaaagacacgacgacgac 960
DB 966 attcattgtcgggtgacaggttggtgctgttcaactctctatctatgaaggtccacgacgac 1025
QY 961 gtgtggaattgtcaagagcttaacgtccatgtctgacgacgacgacgacgacgacgacgacgac 1020
DB 1026 ggttaagattttagtggttccctcaacgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtata 1085
QY 1021 ccattcgttaaggttccggt 1080
DB 1086 ccataagaacggttgacgaggt 1145
QY 1081 tccctaagtgttccatcaatgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 1140
DB 1146 tcaactgacacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 1205
QY 1141 acataagctctccatcaatgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 1200
DB 1206 atgtcgtctcaagtaacatgagaaataaacaacacgacgacgacgacgacgacgacgacgac 1265
QY 1201 agcgacgtgttgtaaaccttaagatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1260
DB 1266 caaaactttagataatcttcaaaaactccgacatgtctcctaggttccagttcaagagc 1325
QY 1261 ttcctgagagcgc 1273
DB 1326 gacctccttgaggc 1338

RESULT 14
AAZ58259
ID AAZ58259 standard; cDNA; 1990 BP.
XX
AC AAZ58259;
XX
XX
DT 08-MAY-2000 (first entry)
XX
DE Rice histone deacetylase gene 1 (HDI) cDNA config.
XX
XX
KW Chromatin associated protein; histone deacetylase gene 1; HD1;
XX rice; transgenic plant; transcription regulation; ss.
XX
OS Oryza sativa.
XX
XX
FT Key Location/Qualifiers
FT CDS 3..1484
FT /tag= a
FT /partial
XX
XX
PD WO200004177-A1.
XX
XX
XX 27-JAN-2000.
XX
XX
XX 13-JUL-1999; 99MO-US15807.
XX
XX 14-JUL-1998; 98US-0092841.
XX
XX
XX (DUP0) DU POINT DE MEMOIRS & CO E. I.

THIS PAGE BLANK (USPTO)